



Génomique et post-génomique du parasite intestinal Blastocystis sp. sous-type 7. Evaluation de son pouvoir pathogène

Ivan Wawrzyniak

► To cite this version:

Ivan Wawrzyniak. Génomique et post-génomique du parasite intestinal Blastocystis sp. sous-type 7. Evaluation de son pouvoir pathogène. Sciences agricoles. Université Blaise Pascal - Clermont-Ferrand II, 2012. Français. NNT : 2012CLF22243 . tel-00741967

HAL Id: tel-00741967

<https://theses.hal.science/tel-00741967>

Submitted on 15 Oct 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Blaise Pascal

N°D.U :2243

**ECOLE DOCTORALE DES SCIENCES DE LA VIE, SANTE,
AGRONOMIE, ENVIRONNEMENT**

N° d'ordre : 575

Thèse

**Présentée à l'Université Blaise Pascal
pour l'obtention du grade de DOCTEUR D'UNIVERSITE**

Spécialité : Microbiologie - Parasitologie

soutenue publiquement le 3 février 2012 par

WAWRZYNIAK Ivan

**Génomique et post-génomique du parasite intestinal
Blastocystis sp. sous-type 7.
Evaluation de son pouvoir pathogène**

Membres du jury :

Président : **Mme BROCHIER-ARMANET Céline**, Professeur, Université Lyon I
Rapporteurs : **Mme FLORENT Isabelle**, MCU, MNHN, Paris
M ALIOUAT El Moukthar, Professeur, Université Lille II
Examineurs : **M VIVARES Christian**, Professeur, Université Clermont-Ferrand II
M EL ALAOUI Hicham, MCU, Université Clermont-Ferrand II
Directeur de thèse : **Mme TEXIER Catherine**, MCU, Université Clermont-Ferrand II

Membre invité :

M DEI CAS Eduardo, MCUPH, Université Lille II

LMGE, UMR-CNRS 6023
Equipe Interactions Hôtes-Parasites
Université Blaise Pascal Clermont-Ferrand II

A mes parents, sans qui je n'en serais pas là aujourd'hui.
A ma femme, Sonia, pour son soutien inconditionnel.
A mon Loulou, pour les sourires dont il me gratifie les matins et les soirs.

Remerciements

Je remercie sincèrement,

Madame Céline Brochier Armanet pour avoir aimablement accepté d'examiner ce travail et de prendre la présidence du jury.

Madame Isabelle Florent et Monsieur El Moukhtar Aliouat pour avoir aimablement accepté d'être les rapporteurs de cette thèse, ainsi que Monsieur Eduardo Dei Cas pour avoir accepté d'examiner ce travail.

Messieurs Christian Vivares et Frédéric Delbac, ancien et nouveau responsable de l'équipe Interactions Hôtes-Parasite (IHP), pour m'avoir fait confiance et permis de réaliser cette thèse. Merci à vous !

Monsieur Christian Amblard, Directeur de l'UMR CNRS 6023 pour m'avoir autorisé à réaliser ce travail.

Madame Catherine Texier et Monsieur Hicham El Alaoui, les directeurs de cette thèse, pour tous nos échanges, vos conseils et votre soutien. Cela a été et demeurera un plaisir de travailler avec vous.

Tous les membres de l'équipe IHP, pour leurs aides techniques, scientifiques et les bons moments passés ensemble.

SOMMAIRE

Liste des abréviations.....	1
Introduction.....	2
Chapitre I, Etude bibliographique.	7
I. <i>Blastocystis</i> spp. un microorganisme parasite méconnu.	8
1. <i>Blastocystis</i> spp. un Straménopile parasite zoonotique à la prévalence sous-estimée.....	8
1.1 <i>Blastocystis</i> spp. un Straménopile méconnu.....	8
1.2 Une grande diversité génétique pour une spéciation compliquée.	10
1.3 Une prévalence longtemps sous-estimée.....	11
1.3.1 Des méthodes de diagnostic de plus en plus performantes.....	11
1.3.1.1 La base du diagnostic : la microscopie optique.....	12
1.3.1.2 Une alternative intéressante : la mise en culture systématique ...	13
1.3.1.3 Le diagnostic moléculaire	13
1.3.1.4 Une sensibilité variable des méthodes de diagnostic.....	13
1.3.2 Une prévalence variable géographiquement.	14
1.4 Une transmission oro-fécale et zoonotique	15
2. <i>Blastocystis</i> spp., un parasite polymorphe au cycle de vie incertain.....	17
2.1 Un organisme à multiples formes	17
2.1.1 La forme vacuolaire.....	18

2.1.2 La forme granulaire.....	19
2.1.3 La forme amiboïde	20
2.1.4 La forme kystique.....	20
2.2 Un cycle de vie encore hypothétique	21
2.2.1 Les différentes formes présentes en culture <i>in vitro</i>	21
2.2.2 Cycle de vie.....	22
3. <i>Blastocystis</i> spp., commensal ou pathogène ?.....	23
3.1 Un organisme associé à des désordres intestinaux chez l'homme	24
3.2 Des modèles d'études <i>in vitro</i> pour évaluer la pathogénie	26
3.3 Des réponses peuvent être apportées par les modèles <i>in vivo</i>	28
3.4.1 Un modèle limité : la souris.....	28
3.4.2 Un modèle plus sensible : le rat.....	29
3.5 <i>Blastocystis</i> spp., un parasite pathogène ?	31
II. La génomique, une avancée dans la connaissance des organismes	32
1. La génomique pour l'étude des parasites.	32
1.1 L'intérêt de séquencer un génome complet.	33
1.2 La reconstruction métabolique.....	34
1.3 La génomique comparative	36
2. Les Straménopiles, classification et génomique	38
2.1 Classification d'un phylum très hétérogène	38
2.2 La génomique des Straménopiles	39

3. Un exemple des intérêts de la génomique chez les parasites intestinaux, le cas des amibes du genre <i>Entamoeba</i>	42
3.1 Caractéristiques et génomiques des espèces du genre <i>Entamoeba</i>	42
3.2 L'analyse du génome d' <i>E. histolytica</i> , un apport de nouvelles données.	43
3.3 <i>E. invadens</i> et <i>E. histolytica</i> , deux amibes pathogènes relativement proches.	44
3.4 <i>Entamoeba dispar</i> , une amibe non pathogène utile pour la génomique comparative.....	45
III. Les protéases, des enzymes essentielles dans la biologie des parasites.....	48
1. Les protéases, un vaste ensemble	48
1.1 Les quatre grandes classes de peptidases	49
1.1.1 Les protéases à sérine	49
1.1.2 Les protéases à cystéine	50
1.1.3 Les métalloprotéases	52
1.1.4 Les protéases à aspartate ou protéases acides.....	54
2. Rôle des protéases chez les protozoaires parasites	56
2.1 Quelques exemples de l'implication des protéases dans les phénomènes de pathogénie	56
2.1.1 Implication dans l'adhésion des parasites aux épithéliums	56
2.1.2 Implication des protéases dans l'invasion tissulaire et cellulaire.....	58
2.1.3 Le dékystement.....	59
2.1.4 Les protéases comme facteurs de virulence.....	60

2.1.4.1 Virulence et facteurs de virulence : définitions	60
2.1.4.2 Les protéases facteurs de virulence.....	61
2.2 Comment les parasites échappent à la réponse immune.....	62
2.2.1 Immunosuppression	62
2.2.2 Immunomodulation de la réponse immune	64
3. Les protéases, des cibles intéressantes dans la lutte contre les parasites ?	66
3.1 Des cibles thérapeutiques.....	66
3.2 Des cibles vaccinales	67
3.3 Des marqueurs pour le séro-diagnostic	68
Chapitre II, Résultats expérimentaux.	70
I. Séquençage des génomes nucléaire et mitochondrial de <i>Blastocystis</i> sp. sous-type 7..	71
II. Prédiction du sécrétome de <i>Blastocystis</i> sp. ST7 et identification de potentiels facteurs de virulence	95
III. Caractérisation de deux cystéines protéases sécrétées par <i>Blastocystis</i> sp ST7.	104
Chapitre III, Conclusions et perspectives.	123
I. Un génome qui a évolué	125
II. Un métabolisme adapté aux conditions d'anoxie du tractus intestinal	127
III. Un réel potentiel pathogène	128
Chapitre IV, Références bibliographiques	132

Liste des abréviations

ATP : Adénosine triphosphate
CoA : Coenzyme A
DAPI : 4',6'-diamidino-2-phénylindole
DTT : Dithiothréitol
EST : Expressed sequence tag
EF : Facteur d'élongation
FECT : technique de concentration au formol éther
GM-CSF : granulocyte-macrophage colony stimulating factor
HSP : Heat shock protein
IBS : Irritable bowel syndrome
Ig : Immunoglobuline
IL : interleukine
KEGG : Kyoto Encyclopedia of Genes and Genomes
Mb : Méga-base
MDR : Multi drug resistance
MET : Microscopie électronique à transmission
MLO : Mitochondrion like organelle
NCBI : National Center for Bioinformatics
NRPS : non ribosomal peptide synthase
PAR : Protease activated receptor
PCR : Polymerase chain reaction
PFOR : Pyruvate ferredoxine oxydoréductase
PKS : Polyketide synthase
PMSF : Phenylmethylsulfonyl fluoride
RFLP : Restriction fragment length polymorphism
RMN : Résonance magnétique nucléaire
SIDA : Syndrome d'immuno-déficience acquise
UNIPROT : Universal Protein Resource
VIH : Virus de l'immuno-déficience humaine

Introduction

Blastocystis spp. est un protozoaire intestinal retrouvé dans le tractus intestinal de l'homme et de divers animaux. Placé dans le groupe hétérogène des Straménopiles, il fut décrit pour la première fois par Alexeieff en 1911, et dénommé *Blastocystis hominis* une année plus tard par Brumpt (Brumpt, 1912), dénomination qui n'est remise en cause que depuis quelques années suite à de multiples études phylogénétiques. Cet organisme, d'abord considéré par la communauté scientifique comme un commensal du tube digestif de l'homme, resta longtemps méconnu. Il faut attendre la fin des années 80 et les travaux de Zierdt (Zierdt, 1978, 1988, 1991) suggérant sa pathogénie pour que ce microorganisme suscite un réel intérêt en tant que parasite. Pourtant ce n'est que depuis la dernière décennie que ce parasite est vraiment étudié, aussi bien au niveau de sa classification phylogénétique, de sa variabilité génétique, de son métabolisme, que de son pouvoir pathogène. Cependant, les études portant sur ce dernier aspect restent discordantes. Certaines l'associe à des symptômes gastro-intestinaux aspécifiques, l'impliquant dans des maladies humaines, d'autres l'exonère, maintenant ainsi la controverse sur son pouvoir pathogène. Ceci trouve une partie de son explication dans la difficulté à identifier le parasite en raison de l'absence de méthodes standardisées de diagnostic, mais également en raison du faible nombre de cas analysés lors d'études épidémiologiques ou encore d'une mauvaise interprétation des données. A cela vient s'ajouter la symptomatologie qui n'est pas systématique chez les patients porteurs de *Blastocystis* spp. et la possibilité que ces symptômes soient dûs à d'autres causes. De plus sa grande diversité génétique a probablement contribué à sous-estimer son importance clinique, de récentes études suggérant une pathogénie variable selon les sous-types (Tan et al., 2010). A l'heure actuelle, 13 sous-types ont été clairement définis sur la base d'une séquence partielle de l'ARNr 18S (Parkar et al., 2010), dont 9 d'entre eux capables d'infecter l'homme (Stensvold et al., 2007a).

Ces dernières années ont vu la communauté scientifique s'intéresser de plus en plus à l'aspect pathogénique de *Blastocystis* spp. De nombreux chercheurs s'accordent sur le fait que *Blastocystis* spp. est un pathogène émergent (Tan et al., 2010). Il est, à ce

jour, le parasite le plus fréquemment rencontré dans les selles de patients humains et présente une prévalence plus élevée que des parasites intestinaux tels que *Giardia* spp., *Entamoeba* spp. ou *Cryptosporidium* spp. (Boorom et al., 2008). Sa prévalence peut atteindre 5% dans les pays développés (Tan, 2008) et 60% dans les pays en développement (Saksirisampant et al., 2006). De plus, des études montrent que l'infection par *Blastocystis* spp. peut être associée à des symptômes gastro-intestinaux aspécifiques (Tan et al., 2010). Le parasite est également associé à des troubles fonctionnels intestinaux chez des patients souffrant d'IBS (Yakoob et al., 2004; Dogruman-Al et al., 2010; Yakoob et al., 2010a). Plusieurs études *in vitro* ont montré que *Blastocystis* sp. est capable de sécréter des protéases pouvant altérer l'épithélium intestinal (Puthia et al., 2005; Puthia et al., 2006; Sio et al., 2006; Mirza et Tan, 2009), et moduler ou dégrader la réponse immunitaire (Long et al., 2001; Puthia et al., 2005; Puthia et al., 2008). *Blastocystis* sp. apparaît donc comme un parasite susceptible d'agir de manière directe ou indirecte sur son hôte et ainsi d'engendrer des symptômes.

Dans le cadre des recherches sur ce parasite, l'objectif de l'équipe IHP était d'étudier la pathogénie et le métabolisme de ce parasite anaérobie, en utilisant les données issues du séquençage du génome du sous-type 7 de *Blastocystis* sp. Ma thèse s'est inscrite dans le cadre de cet objectif. J'ai eu donc en charge la mise en culture du parasite, la production de l'ADN utilisée pour le séquençage, et la mise en place d'une liste de gènes permettant l'annotation automatique du génome. J'ai également participé à l'analyse du génome avec pour objectif l'identification de gènes pouvant coder des protéines potentiellement impliquées dans la pathogénie. Les analyses *in silico* ont montré la présence de nombreux gènes candidats, et notamment des gènes codant des protéases. A partir des données issues de cette analyse, j'ai orienté ma recherche vers les protéases prédites sécrétées à travers la caractérisation *in vitro* d'activités protéasiques dans les surnageants de culture, permettant ainsi l'identification de protéines candidates pouvant induire des symptômes intestinaux.

Mon travail de thèse a donné lieu aux publications suivantes :

• Articles :

Poirier P, **Wawrzyniak I**, Vivarès CP, Delbac F, and El Alaoui H.

New insights into *Blastocystis* spp.: a potential link with the irritable bowel syndrome. PLoS Pathog 8(3): e1002545, doi:10.1371.

Wawrzyniak I, Texier C, Poirier P, Viscogliosi E, Tan KSW, Delbac F and El Alaoui H.

Characterization of two cysteine proteases secreted by *Blastocystis* sp., a human intestinal parasite. Parasitol Int (2012), doi: 10.1016.

Denoeud F, Roussel M, Noel B, **Wawrzyniak I**, Da Silva C, Diogon M, Viscogliosi E, Brochier-Armanet C, Couloux A, Poulain J, Segurans B, Anthouard V, Texier C, Blot N, Poirier P, Geok Choo N, Tan KSW, Artiguenave F, Jaillon O, Aury JM, Delbac F, Wincker P, Vivarès CP, El Alaoui H.

Genome sequence of the stramenopile *Blastocystis*, a human anaerobic parasite. Genome Biol, 2011. 12:R29.

Wawrzyniak I, Roussel M, Diogon M, Couloux A, Texier C, Tan K S, Vivares C P, Delbac F, Wincker P, El Alaoui H.

Complete circular DNA in the mitochondria-like organelles of *Blastocystis hominis*. Int J Parasitol, 2008. 38(12): 1377-82.

• Communications orales :

Wawrzyniak I, Poirier P, Roussel M, Delbac F, Texier C, El Alaoui H.

Post-génomique du parasite *Blastocystis* sp : identification de cystéine protéases secrétées et potentiellement impliquées dans la pathogénie. Congrès Groupement des protistologues de Langue Française (GPLF) 2010, Tunis, Tunisie.

Wawrzyniak I, El Alaoui H, Chambon C, Poirier P, Roussel M, Delbac F, Texier C.

Etude génomique du parasite intestinal *Blastocystis hominis*. Application à sa pathogénie. Journées des écoles doctorales, Clermont-Ferrand 2010.

L'utilisation des données issues de l'analyse génomique que j'ai réalisée m'a permis de participer aux publications suivantes :

Poirier P, **Wawrzyniak I**, Albert A, El Alaoui H, Delbac F, Livrelli V.

Development and evaluation of a Real-Time PCR assay for detection and quantification of *Blastocystis* parasites in human stool samples: prospective study of patients with Hematological Malignancies. J Clin Microbiol, 2011. 975-983.

Souppart L, Sanciuciu G, Cian A, **Wawrzyniak I**, Delbac F, Capron M, Dei-Cas E, Boorom K, Delhaes L, Viscogliosi E.

Molecular epidemiology of human *Blastocystis* isolates in France. Parasitol Res, 2009. 105(2): 413-421.

Avant la réalisation de ma thèse, qui a débuté en novembre 2007, j'ai été recruté en 2002 au Laboratoire Microorganismes : Génome et Environnement (ex Laboratoire de Biologie des Protistes) en tant qu'assistant-ingénieur CNRS. Ma collaboration aux travaux de l'équipe IHP m'a également permis de participer aux publications suivantes :

Polonais V, Mazet M, **Wawrzyniak I**, Texier C, Blot N, El Alaoui H, Delbac F.

The human microsporidian *Encephalitozoon hellem* synthesizes two spore wall polymorphic proteins useful for epidemiological studies. Infect Immun, 2010, 78(5): 2221-2230.

Chavant P, Taupin V, El Alaoui H, **Wawrzyniak I**, Chambon C, Prensier G, Metenier G, Vivares C P.

Proteolytic activity in *Encephalitozoon cuniculi* sporogonial stages: predominance of metallopeptidases including an aminopeptidase-P-like enzyme, Int J Parasitol, 2005. 35(13): 1425-1433.

Chapitre I.

Etude bibliographique

Tableau 1 : Caractéristiques morphologiques et physiologiques utilisées par Zierdt pour classer *Blastocystis* sp. parmi les protistes (Zierdt, 1991).

Characteristic	Yeast	<i>B. hominis</i> *
Strict anaerobe	No	Yes
No growth on fungal or bacteriological media	No	Yes
Growth on media developed for intestinal protozoa	No	Yes
No growth on surface of solid media	No	Yes
Requires presence of bacteria for growth; axenic growth achieved slowly under carefully controlled conditions	No	Yes
Capable of ingesting bacteria and other particulate matter	No	Yes
Cultures die in 2–3 days at 22°C or overnight at 4°C	No	Yes
No growth below 33°C, but cell death at 30°C	No	Yes
Optimal growth at 37°C	No	Yes
Optimal growth at neutral pH	No	Yes
No growth at pH 5.5	No	Yes
Resistant to 400 µg of amphotericin per ml	No	Yes
Sensitive to drugs effective against intestinal protozoa	No	Yes
No cell wall	No	Yes
Reproduction by endodyogony, schizogony, binary fission, and plasmotomy (cutting off) individual <i>B. hominis</i> from the ameba form	No	Yes
No budding	No	Yes
Supports stable bacterial endosymbiont (an obligate mutualism)	No	Yes
Slow-feeding pseudopods	No	Yes
Rapid locomotion pseudopods	No	Yes
Limiting membrane with micropinocytotic vesicles	No	Yes
Membrane-bound CB, a reproductive organelle, formerly called vacuole	No	Yes
Mitochondria in all cells; mitochondria show a general morphology similar to that in other protozoa, with short saccate cristae when in the resting mode	No	Yes

* *Blastocystis* spp. ne présente pas les mêmes propriétés que les levures.

I. *Blastocystis* spp. un microorganisme parasite méconnu.

Blastocystis spp. est un protozoaire parasite intestinal de l'homme et des animaux. Décrit dans les années 1900, son caractère pathogène reste controversé notamment en raison de l'existence de porteurs asymptomatiques (Tan, 2008). Pourtant, il est le parasite le plus identifié dans les selles de patients humains (Boorom et al., 2008) et est susceptible d'entraîner de multiples symptômes gastro-intestinaux aspécifiques comme des diarrhées, des vomissements, des nausées, des douleurs abdominales (Tan, 2008). De plus, des études suggèrent son implication dans des cas de syndrome de l'intestin irritable (Dogruman-Al et al., 2010). Sa répartition géographique est mondiale, avec une prévalence variable selon les pays, de 5% dans les pays développés (Tan, 2008) jusqu'à 60% dans les pays en développement (Pegelow et al., 1997). Cette prévalence est certainement liée au mode de transmission oro-fécale du parasite et aux conditions d'hygiène très disparates selon les pays.

L'intérêt croissant porté à ce parasite s'est traduit par une description morphologique plus poussée, la révision de sa classification, l'évaluation de sa diversité génétique, l'évaluation de son pouvoir zoonotique, la mise en place de nouvelles méthodes de diagnostic ou le développement de modèles d'interactions hôte-parasite. Cependant, plusieurs interrogations subsistent, comme la structure et la fonction des organites apparentés aux mitochondries, son pouvoir pathogène et les mécanismes associés, son cycle de vie, ou encore la nécessité d'un traitement.

1. *Blastocystis* spp. un Straménopile parasite zoonotique à la prévalence sous-estimée

1.1 *Blastocystis* spp. un Straménopile méconnu

La taxinomie de *Blastocystis* spp. est restée longtemps contestée, et les multiples changements de classification dont il fut l'objet depuis les années 1970, reflètent les difficultés à lui trouver une position taxinomique fixe. En 1967, sur la base d'études

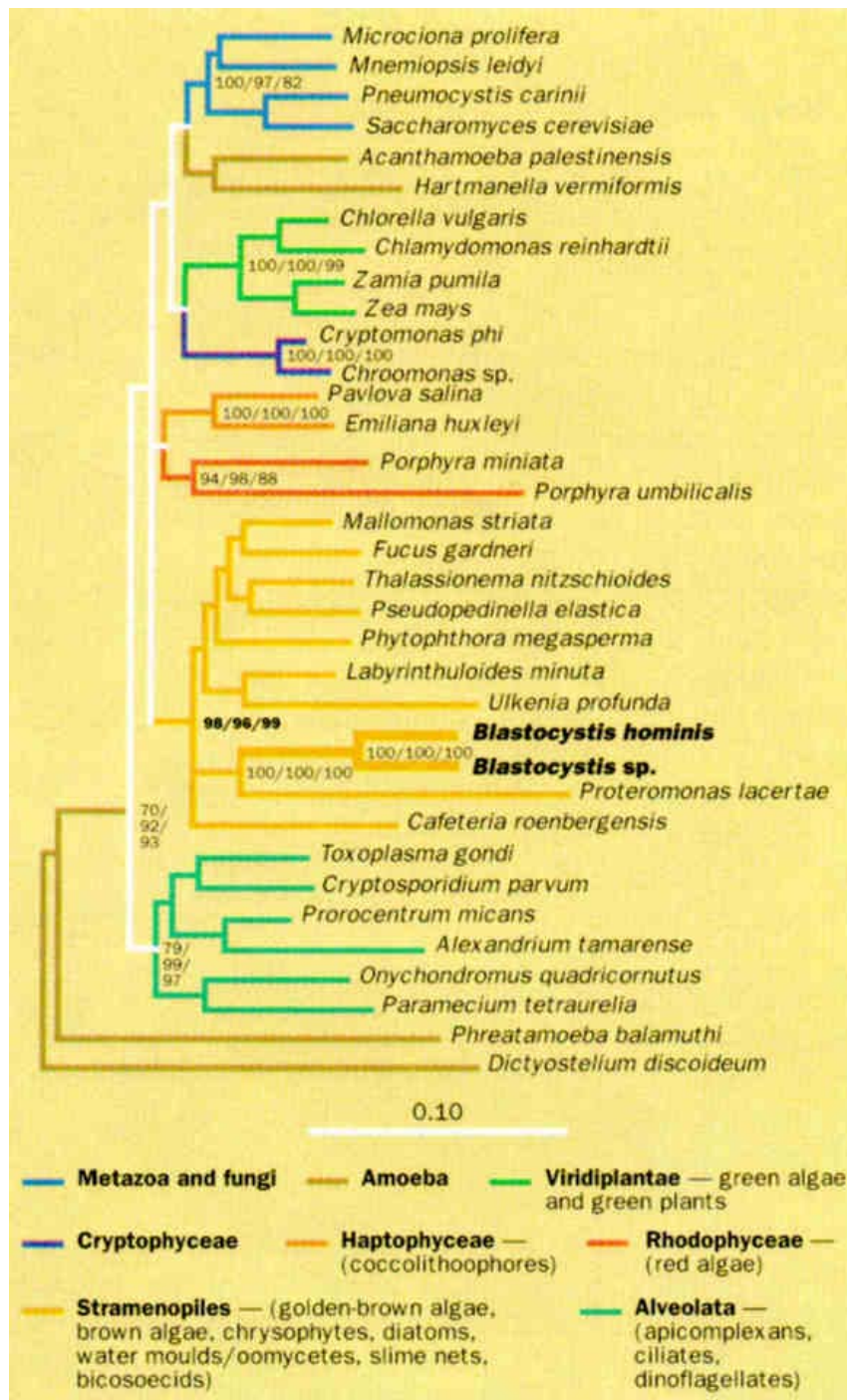


Figure 1: Classification taxinomique de *Blastocystis* spp. réalisée à partir de la séquence complète de l'ARNr 18S (Silberman et al., 1996). Les deux isolats (un isolat humain et un isolat issu du cochon d'inde) sont très proches l'un de l'autre, les séquences d'ARNr 18s ne présentent que 6,4 % de différence d'homologie. Les valeurs élevées de bootstrap et l'analyse des distances placent *Blastocystis* sp. parmi les Straménopiles.

morphologiques et physiologiques (Tableau 1), il fut classé parmi les protistes (Zierdt et al., 1967), alors qu'il était depuis sa découverte décrit comme une levure ou un champignon, en raison de son apparence proche des champignons et de son mode de division par fission binaire. En 1978, Zierdt (Zierdt, 1978) le classa dans le sous-phylum des Sporozoaires, en se basant sur son mode de culture, ainsi que sur de nouvelles observations morphologiques. Finalement, en raison de sa localisation exclusivement extracellulaire et de l'absence de pôle apical, il fut classé dans les Sarcodinés¹ en 1988 (Zierdt et al., 1988; Zierdt, 1991).

L'arrivée des premières analyses moléculaires, basées sur le séquençage partiel de l'ARNr de la petite sous unité ribosomique (ARNr18S), a permis de montrer que *Blastocystis* spp. n'était pas monophylétique avec les levures, les champignons ou les Sarcodinés (Johnson et al., 1989). Il fut donc proposé en 1993 un nouveau sous-phylum : *Blastocysta*, incluant la classe *Blastocystea*, l'ordre *Blastocystida*, la famille *Blastocystidae* et le genre *Blastocystis* avec *Blastocystis hominis* comme espèce (Jiang et He, 1993). L'étude de la totalité de la séquence de l'ARNr18S par Silberman (Silberman et al., 1996) montra qu'il pouvait être placé parmi les Straménopiles. Ce groupe complexe très hétérogène regroupe à la fois des organismes unicellulaires et pluricellulaires, hétérotrophes ou photosynthétiques pouvant aller des parasites végétaux du genre *Phytophthora* spp. aux algues brunes en passant par les diatomées. Le phylum des Straménopiles fut créé en 1986 par Cavalier-Smith (Cavalier-Smith, 1986) et regroupait initialement des organismes caractérisés par la présence au cours de leur cycle de développement d'un stade qui possède un flagelle à mastigonème². Silberman révèle que *Blastocystis* spp. est proche phylogénétiquement du flagellé *Proteromonas lacertae* (Figure 1). Ces deux organismes partagent certaines caractéristiques, notamment le fait d'être des parasites de l'intestin de vertébrés ou encore de posséder un stade de résistance qui est aussi le stade de contamination. Cependant au contraire de

¹ Les Sarcodinés sont un phylum de protozoaires comprenant plus de 11000 espèces. Il inclut notamment les amibes. Ils sont caractérisés par l'utilisation de pseudopodes pour se déplacer ou pour la capture d'aliments. La plupart sont des organismes libres, même si certains sont des parasites.

² Les mastigonèmes sont des filaments situés sur la membrane de certains flagelles

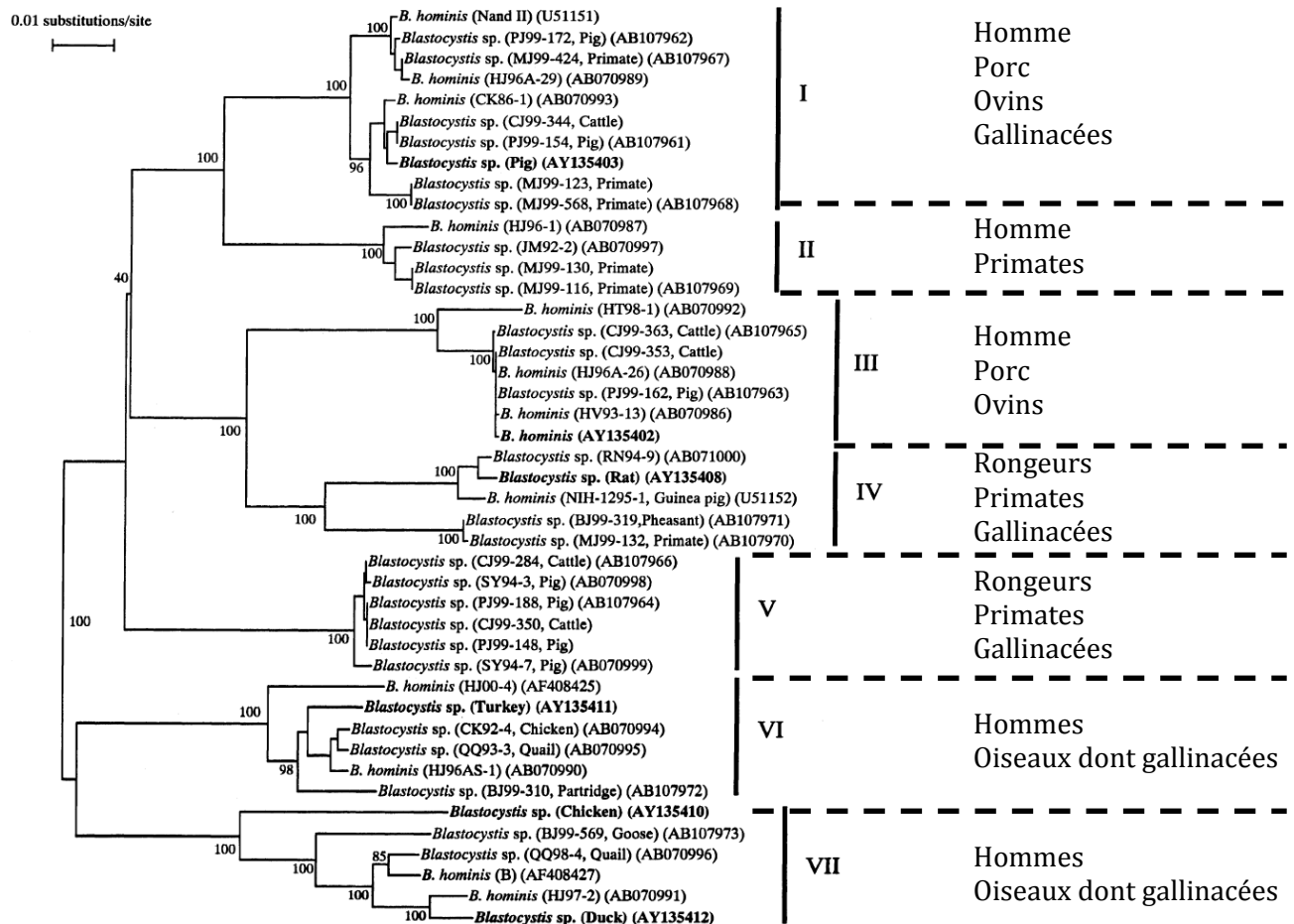


Figure 2: Analyse phylogénétique de différents isolats humains et animaux de *Blastocystis* spp. basé sur la séquence complète de l'ARNr 18S par la méthode de Neighbour joining (Abe et al, 2004). Les séquences des ARNr 18S qui ne sont que partielles apparaissent en gras. Le groupe I consiste en isolats humains, porcins, ovins et galliformes, le groupe II en isolats humains et de primates, le groupe III en isolats humains, porcins et ovins, le groupe IV en isolats murins, en un isolat de primate et en un isolat de faisan, le groupe V en isolats porcins et ovins, et les groupes VI et VII en isolats humains et en isolats aviaires.

Proteromonas lacertae, *Blastocystis* spp. ne possède pas de flagelle à mastigonème. Une étude basée sur l'analyse phylogénique moléculaire de plusieurs gènes : l'ARNr 18S, une HSP 70, le facteur d'élongation EF2, et la sous-unité 2 de l'ATPase vacuolaire, vient confirmer la classification du parasite parmi les Straménopiles (Arisue et al., 2002).

1.2 Une grande diversité génétique pour une spéciation compliquée.

Blastocystis spp. a été isolé à partir d'échantillons issus d'une multitude d'espèces incluant les primates, les cochons, les rongeurs, les reptiles ou encore les oiseaux (Stenzel et Boreham, 1996). Aucune différence morphologique significative n'ayant été mise en évidence entre les isolats, les critères morphologiques n'ont pas pu servir pour étudier la spéciation du parasite. Plusieurs méthodes ont alors été testées pour essayer de différencier les isolats (Tan, 2004): la RFLP, la PCR avec des amorces spécifiques de sous-type et la variabilité caryotypique. Toutes ces méthodes ont mis en évidence une forte variabilité génétique. Ainsi, La PCR-RFLP, utilisée sur l'ARNr 18S de 29 isolats de *Blastocystis* spp. (Clark, 1997) a mis en évidence l'existence de 7 profils de restriction différents.

Quelques années plus tard, l'analyse de l'ARNr 18S sur 16 isolats humains et animaux de *Blastocystis* spp. a montré que l'on pouvait séparer les isolats en 7 groupes phylogénétiques distincts, ayant une morphologie identique mais étant génétiquement différents (Arisue et al., 2003). Parmi ces 7 groupes phylogénétiques, 5 regroupent à la fois des isolats humains et animaux. Les auteurs suggèrent que ces 7 groupes correspondent à des espèces différentes. Une autre analyse de l'ARNr 18S de 44 isolats différents, humains et animaux, a permis de les classer en 7 groupes monophylétiques (Abe, 2004) correspondant exactement aux 7 groupes décrits par Arisue (Figure 2). Cette étude met également en évidence la faible spécificité d'hôtes et le fort potentiel zoonotique des différents isolats de *Blastocystis* spp.

En 2005, Noel et al, à l'aide d'une étude phylogénétique regroupant les 78 séquences de *Blastocystis* spp. présentes dans les bases de données et en y ajoutant 12

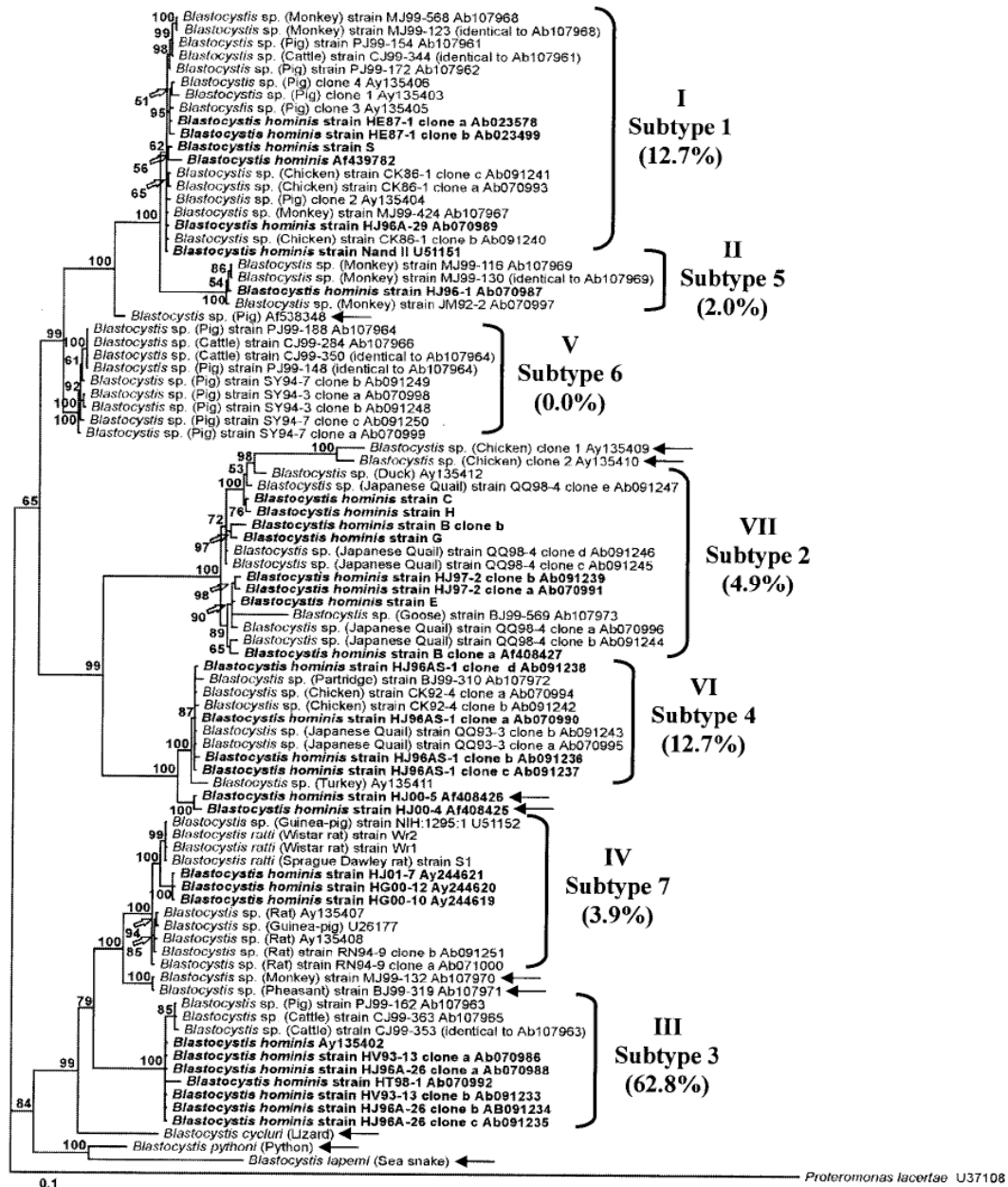


Figure 3 : Analyse phylogénétique des différents isolats de *Blastocystis* spp. établie par la méthode du maximum de vraisemblance à partir de la séquence du gène codant l'ARNr 18S (Noel et al, 2005). *Proteromonas lacertae* est utilisé comme groupe externe. Les valeurs en pourcentage correspondent aux probabilités postérieures bayésiennes. La probabilité postérieure d'un arbre peut être interprétée comme la probabilité que cet arbre soit vrai connaissant les données. Les isolats humains sont notés en caractère gras. Les groupes I à VII correspondent à la classification précédemment proposée par Arisue et al (2003). Les flèches indiquent les isolats pour lesquels la classification est incertaine.

séquences d'ARNr 18S d'isolats de *Blastocystis* sp. humains, murins et de reptiles obtenus après amplification par PCR, confirma l'existence de 7 groupes majeurs référencés comme les sous-types 1 à 7 (Noel et al., 2005). Une corrélation fut alors établie entre les sous-types identifiés et les groupes issus des analyses phylogénétiques précédentes (Arisue et al., 2003; Abe, 2004; Noel et al., 2005) (Figure 3).

Finalement les différentes études ont abouti à une certaine confusion sur la terminologie permettant l'identification et la désignation des différents isolats. Un consensus a donc été établi pour typer les isolats de *Blastocystis* spp. (Stensvold et al., 2007b). Les différents isolats étant désormais désignés comme *Blastocystis* sp. avec un numéro de sous-type allant de ST1 à ST9 (Tableau 2). L'homme est supposé être l'hôte spécifique du ST3, les autres STs étant retrouvés chez une large variété d'animaux incluant les primates, les ovins, les bovins, les porcins, les chiens (ST 1, 2, 5, 8), les rongeurs (ST 4) et les oiseaux (ST 6, 7, 8). Le ST 9 consiste actuellement en 2 isolats retrouvés chez l'homme. Par la suite, de nouveaux sous-types ont été identifiés uniquement chez certains animaux, le ST 10 chez des primates (Stensvold et al., 2009), le ST 11 chez les éléphants, le ST 12 chez la girafe, et le ST 13 chez le quokka³ et le kangourou gris (Parkar et al., 2010).

1.3 Une prévalence longtemps sous-estimée

1.3.1 Des méthodes de diagnostic de plus en plus performantes

En raison de sa pathogénie controversée, *Blastocystis* spp. reste peu recherché lors des analyses médicales, les praticiens ne le considérant pas comme un agent étiologique responsable de maladie. De plus, son polymorphisme rend son diagnostic difficile en laboratoire et nécessite un technicien expérimenté pour pouvoir le détecter en observation directe dans les échantillons de selles. Cette difficulté d'identification a conduit les biologistes à améliorer les méthodes d'identification et à compléter les

³ Le Quokka est un petit marsupial de la famille des macropidés (le plus connu des macropidés étant le kangourou)

Tableau 2 : Corrélation entre les différentes classifications existantes des sous-types de *Blastocystis* sp. et le consensus proposé par Stensvold et al (Stensvold et al, 2007b).

Clade ^a	Subtype ^b	Group and subtype ^c	Subtype ^d	Ribodème ^{e,f}	Subgroup ^g	Cluster ^h	Subtype ⁱ	Consensus
I	I	I/1	1	1, 8 ^l	III	E	1, 1 variant	<i>Blastocystis</i> sp. subtype 1
II	II	II/5	5	6	V	C, D	– ^k	<i>Blastocystis</i> sp. subtype 2
–	X	I+II/1+5 outlier	–	–	–	–	–	Chimaeric sequence
III	III	III/3	3	2, 7, 4? ^l , 5?	I, II	A	3	<i>Blastocystis</i> sp. subtype 3
IV	IV	IV/7	7	3	IV	B	–	<i>Blastocystis</i> sp. subtype 4
–	IVa	IV/7 outliers	–	–	–	–	–	<i>Blastocystis</i> sp. subtype 8
V	V	V/6	6	–	–	–	–	<i>Blastocystis</i> sp. subtype 5
VI	VI	VI/4	4	9 ^l	–	–	4	<i>Blastocystis</i> sp. subtype 6
–	VIa	VI/4 outliers	–	–	–	–	–	<i>Blastocystis</i> sp. subtype 9
VII	VII	VII/2	2	10	VI ^m	–	2	<i>Blastocystis</i> sp. subtype 7
–	VII	VII/2 outliers	–	–	–	–	–	<i>Blastocystis</i> sp. subtype 7

a : Clades décrits par Arisue *et al.* (2003) et Yoshikawa *et al.* (2004).

b : Sous-types décrits par Scicluna *et al.* (2006).

c : Groupes et sous-types décrits par Noel *et al.* (2005).

d : Sous-types décrits par Yoshikawa *et al.* (1998, 2000).

e,f : Ribodèmes décrits par Clark (1997), et Yoshikawa *et al.* (2000).

g : Sous-groupes décrits par Böhm-Gloning *et al.* (1997).

h : Clusters décrits par Stensvold *et al.* (2006).

i : Sous-types décrits par Yoshikawa *et al.* (2000).

analyses directes en microscopie optique par la mise en culture systématique ou par l'utilisation de méthodes moléculaires.

Le choix et la fiabilité des méthodes de diagnostic ont directement influencé les études de prévalence de *Blastocystis* spp. Cette influence a eu tendance à sous-estimer la prévalence du parasite. Le développement de méthodes de diagnostic par PCR a permis d'éliminer ce biais, et a eu l'avantage de permettre de typer les différents isolats cliniques.

1.3.1.1 *La base du diagnostic : la microscopie optique*

Le diagnostic en microscopie optique à partir d'un examen direct des selles des patients est la méthode la plus commune pour détecter *Blastocystis* spp. Cependant, compte tenu des multiples formes existantes du parasite, de leurs altérations provoquées par des conditions environnementales ou des traitements médicamenteux, du risque de confondre *Blastocystis* spp. avec d'autres microorganismes, cette méthode semble avoir entraîné une sous-estimation de la prévalence du parasite (Leelayoova et al., 2002; Suresh et Smith, 2004; Termmathurapoj et al., 2004). De plus, la forme classiquement détectée est la forme vacuolaire, forme minoritaire dans les échantillons frais de selles, en comparaison des formes kystiques largement majoritaires mais difficilement identifiables (Tan, 2004). La détection des formes kystiques du parasite nécessite des étapes de concentration (Suresh et Smith, 2004), même si la méthode de concentration la plus utilisée, la FECT (formol ether concentration technique), n'est pas recommandée pour l'identification de *Blastocystis* spp. en raison d'une faible sensibilité (Tan et al., 2010). Les colorations couramment utilisées pour les observations directes sont le lugol iodine, le Giemsa et le trichrome, ce dernier étant le plus utilisé (Tan, 2008).

1.3.1.2 Une alternative intéressante : la mise en culture systématique

La mise en culture systématique des échantillons s'est avérée une méthode de choix pour confirmer les diagnostics microscopiques. Le milieu de Jones pour la mise en culture xénique *in vitro* à partir d'échantillons de selles s'est avéré le plus efficace (Suresh et Smith, 2004). Cette méthode est jusqu'à six fois plus sensible que l'observation directe en microscopie (Termmathurapoj et al., 2004). Cependant cette méthode présente quelques inconvénients dont le principal est la difficulté voire l'impossibilité à cultiver certains isolats en milieu de Jones (Parkar et al., 2007; Tan, 2008). De plus, le temps de multiplication variable selon les sous types, entraîne la disparition des cas de co-infections lors de la mise en culture (Tan et al., 2010). L'utilisation des méthodes de diagnostic moléculaire s'avère être une meilleure alternative.

1.3.1.3 Le diagnostic moléculaire

Plusieurs approches de diagnostic moléculaire, principalement basées sur un diagnostic PCR ont été décrites ces dernières années. L'utilisation d'amorces spécifiques ciblant l'ARNr 18S a été privilégiée, permettant ainsi le génotypage des sous-types (Yoshikawa et al., 1998; Stensvold et al., 2007b; Santin et al., 2011). Récemment, deux méthodes de PCR en temps réel ont été développées ciblant respectivement une région de 152 bases d'un gène inconnu (Jones et al., 2008) et une région de l'ARNr 18S (Poirier et al., 2011). L'utilisation de ces méthodes moléculaires a permis la mise en évidence de nouveaux sous-types, mais également la description de co-infections et la réévaluation de la prévalence du parasite.

1.3.1.4 Une sensibilité variable des méthodes de diagnostic

De nombreux travaux ont été menés pour comparer les différentes méthodes de diagnostic existantes. La majorité de ces travaux arrive à la conclusion que les méthodes moléculaires sont les plus sensibles. Certains chiffres apparaissent même éloquentes,

puisque une étude réalisée sur 900 échantillons de selles révèlent 142 échantillons positifs en observation microscopique alors que 334 le sont après mise en culture (Leelayoova et al., 2002). Cette sensibilité supérieure de la culture xénique *in vitro* par rapport à la microscopie optique, a été confirmée par de nombreuses autres études (Surolia et Surolia, 2001; Termmathurapoj et al., 2004; Stensvold et al., 2007a; Poirier et al., 2011). L'utilisation des méthodes moléculaires a permis une détection encore plus sensible du parasite. Stensvold et al (Stensvold et al., 2007a) ont montré que sur 100% d'échantillons positifs en PCR, seuls 89% étaient positifs en culture xénique *in vitro*. En 2010, une étude réalisée au CHU de Clermont-Ferrand sur une cohorte de 186 patients, a révélé 27 patients positifs par PCR contre 14 après culture xénique *in vitro*, et 8 après observation directe au microscope (Poirier et al., 2011).

Sur la base de ces différentes études, le diagnostic de *Blastocystis* spp. en microscopie optique se révèle insuffisant. Il nécessite une mise en culture des échantillons de selles se révélant négatifs. Le diagnostic PCR s'il se révèle plus sensible et fiable que la mise en culture, présente aussi l'intérêt de permettre un sous-typage des différentes souches.

1.3.2 Une prévalence variable géographiquement

Blastocystis spp. possède une répartition mondiale, avec une prévalence variable selon les pays. En effet, celle-ci est plus élevée dans les pays en développement que dans les pays développés, ce constat étant certainement lié aux conditions d'hygiène et/ou à un contact plus étroit avec les réservoirs animaux. Ainsi la prévalence est de l'ordre de 1% au Japon (Hirata et al., 2007) ou de 3,3% à Singapour (Wong et al., 2008) alors qu'elle est de 33,3% en Egypte (Rayan et al., 2007) et de 60% en Indonésie (Pegelow et al., 1997).

Les méthodes moléculaires ont permis de mieux comprendre la répartition des différents sous-types. La majorité des études montre le sous-type 3 comme étant l'isolat le plus commun dans de nombreux pays. Ainsi la prévalence du sous-type 3 est de 53,5%

Tableau 3 : Répartition géographique des différents sous-types d'isolats humains de *Blastocystis* spp.

References	Countries	No. of isolates	Subtype classification ^b (% in parentheses)									
			1	2	3	4	5	6	7	8	9	Unknown ^c
Yoshikawa et al. 2004	Bangladesh	26	2 (7.7)	0	24 (92.3)	0	0	0	0	0	0	0
Li et al. 2007a	China	202 ^a	56 (27.7)	11 (5.4)	125 (61.8)	1 (0.5)	0	1 (0.5)	0	0	0	8 (4.0)
Yan et al. 2006	China	40 ^a	18 (45.0)	0	19 (47.5)	0	0	0	2 (5.0)	0	0	1 (2.5)
Li et al. 2007b	China	80 ^a	18 (22.5)	2 (2.5)	56 (70.0)	1 (1.2)	0	0	0	0	0	3 (3.8)
Stensvold et al. 2007a	Denmark	28	5 (17.9)	9 (32.1)	13 (46.4)	1 (3.6)	0	0	0	0	0	0
Stensvold et al. 2006	Denmark	29	1 (3.4)	6 (20.7)	15 (51.7)	7 (24.1)	0	0	0	0	0	0
Hussein et al. 2008	Egypt	44	8 (18.2)	0	24 (54.5)	0	0	8 (18.2)	4 (9.1)	0	0	0
Clark 1997	England	29	2 (6.9)	1 (3.4)	22 (75.9)	4 (13.8)	0	0	0	0	0	0
Scicluna et al. 2006	England	55	3 (5.4)	9 (16.4)	22 (40.0)	17 (31.0)	0	0	1 (1.8)	3 (5.4)	0	0
Present study	France	43 ^a	11 (25.6)	4 (9.3)	23 (53.5)	4 (9.3)	0	0	1 (2.3)	0	0	0
Yoshikawa et al. 2004	Germany	12	3 (25.0)	2 (16.7)	5 (41.7)	2 (16.7)	0	0	0	0	0	0
Böhm-Gloning et al. 1997	Germany	171 ^a	40 (23.4)	1 (0.6)	118 (69.0)	12 (7.0)	0	0	0	0	0	0
Menounos et al. 2007	Greece	45	9 (20.0)	6 (13.3)	27 (60.0)	1 (2.2)	0	1 (2.2)	1 (2.2)	0	0	0
Kaneda et al. 2001	Japan	64	10 (15.6)	13 (20.3)	31 (48.4)	7 (10.9)	0	0	0	0	0	3 (4.7)
Yoshikawa et al. 2004	Japan	50	4 (8.0)	0	26 (52.0)	2 (4.0)	0	11 (22.0)	5 (10.0)	0	2 (4.0)	0
Yoshikawa et al. 2004	Pakistan	10	2 (20.0)	0	7 (70.0)	0	0	1 (10.0)	0	0	0	0
Wong et al. 2008	Singapore	9	2 (22.2)	0	7 (77.8)	0	0	0	0	0	0	0
Dogruman-Al et al. 2008	Turkey	96 ^a	19 (19.8)	22 (22.9)	55 (57.3)	0	0	0	0	0	0	0
Özyurt et al. 2008	Turkey	87	8 (9.2)	12 (13.8)	66 (75.9)	1 (1.1)	0	0	0	0	0	0
Total		1,120	221 (19.7)	98 (8.7)	685 (61.2)	60 (5.4)	0	22 (2.0)	14 (1.2)	3 (0.3)	2 (0.2)	15 (1.3)

Le sous-type 3 apparaît systématiquement majoritaire (Souppart et al., 2009).

a : inclus les isolats avec des infections multiples

b : selon la nomenclature définie par Stensvold (Stensvold et al., 2007b)

c : Ces isolats n'ont pas été amplifiés avec les amorces standards utilisées dans ces études ; par conséquent, ils n'ont pas été assignés à l'un des sous-types.

en France (Souppart et al., 2009), 61,9 % en Egypte (Souppart et al., 2010), 75,9% en Turquie (Ozyurt et al., 2008) ou 78% à Singapour (Wong et al., 2008). Il en est de même pour le Japon, le Pakistan, le Bangladesh ou encore l'Allemagne (Yoshikawa et al., 2004a). Ces résultats laissent supposer l'absence de corrélation entre un sous-type et une localisation géographique. La prédominance du sous-type 3 peut être expliquée du fait de sa spécificité supposée pour l'homme, et donc par une transmission interhumaine favorisée. La prévalence des autres sous-types est variable même si les sous-types 1, 2 et 4 sont les plus fréquemment rencontrés après le ST 3 (Tableau 3). Cependant, certaines régions ont une prévalence du ST 4 plus élevée que le ST 3. En effet, une étude menée dans le district de Valence en Espagne a montré une prévalence du ST 4 de 94,1% (Dominguez-Marquez et al., 2009). De même des travaux menés sur des patients du CHU de Clermont-Ferrand (Poirier et al., 2011) ont montré une prévalence de 63% du ST 4. Ces distributions variables suggèrent des réservoirs différents.

1.4 Une transmission oro-fécale et zoonotique

Il est actuellement admis que le mode de contamination par *Blastocystis* spp. se fait de manière oro-fécale (Yoshikawa et al., 2004b). La présence d'une forme kystique dans le cycle du parasite, qui est capable de résister dans l'eau et les selles (Moe et al., 1996), laisse supposer que la contamination est liée à de mauvaises conditions d'hygiène, notamment à travers la consommation d'eau souillée et le contact avec des animaux réservoirs du parasite (Garavelli et Scaglione, 1989; Doyle et al., 1990; Nimri et Batchoun, 1994; Leelayoova et al., 2004; Leelayoova et al., 2008; Eroglu et Koltas, 2010; Banaticla et Rivera, 2011).

Des analyses menées sur du personnel militaire thaïlandais ont montré *Blastocystis* spp. comme étant le parasite intestinal prédominant, avec une prévalence allant de 22,4% à 36,4% (Leelayoova et al., 2004; Leelayoova et al., 2009). Cette prévalence a été reliée à une consommation d'eau non filtrée ou non bouillie. L'absence de structure de traitement des eaux usées apparaît comme une des raisons majeures de

contamination des populations (Nimri et Batchoun, 1994; Eroglu and Koltas, 2010; Banaticla et Rivera, 2011). Une analyse de l'eau distribuée dans une école thaïlandaise montre la présence d'un sous-type 1 de *Blastocystis* sp. qui est retrouvé chez 18,9% des enfants (Leelayoova et al., 2008).

L'exposition aux animaux est une autre cause majeure de contamination des populations en raison de l'apparente faible spécificité d'hôte du parasite. Ainsi, *Blastocystis* spp. se révèle comme un parasite très commun au sein du règne animal. Il a été rapporté chez les mammifères, les oiseaux, les reptiles, les amphibiens et les arthropodes (Tan, 2004). Certains animaux présentent même une prévalence très élevée. Parmi les animaux de ferme par exemple, on peut noter une prévalence allant jusqu'à 95 % chez le cochon et 71 % chez le mouton (Abe et al., 2002). Il en est de même pour les animaux domestiques où une prévalence a été observée de 16,8 à 67% chez le chat et de 0 à 70% chez le chien (Duda et al., 1998; Abe et al., 2002; Mohsen et Hossein, 2009). Cette variabilité peut être facilement expliquée par les méthodes de diagnostic utilisées ou des situations épidémiologiques différentes. Les animaux de zoo montrent aussi une forte prévalence. Ainsi, 85% des primates, 80% des faisans et 56% des canards au zoo d'Osaka se révèlent positifs à *Blastocystis* spp. (Abe et al., 2002). L'étude de 76 échantillons de selles issues du zoo de Perth en Australie révèle 42% d'animaux positifs. Ce chiffre est encore plus élevé pour les primates qui se révèlent positifs à 50 % (Parkar et al., 2010).

La contamination zoonotique due à la proximité d'animaux a largement été démontrée (Garavelli et Scaglione, 1989; Doyle et al., 1990), et confirmée par des analyses phylogénétiques désignant ainsi la blastocystose comme une zoonose (Yoshikawa et al., 2004a; Noel et al., 2005). Une étude de prévalence sur les animaux et leurs gardiens, mené au parc zoologique de Perth en Australie, a ainsi mis en évidence une forte contamination des gardiens (63%) en comparaison avec une population contrôle où l'on retrouve seulement 9% de positifs (Parkar et al., 2010). Chez les animaux, 50% des échantillons issus de primates se sont révélés positifs, et 32% des

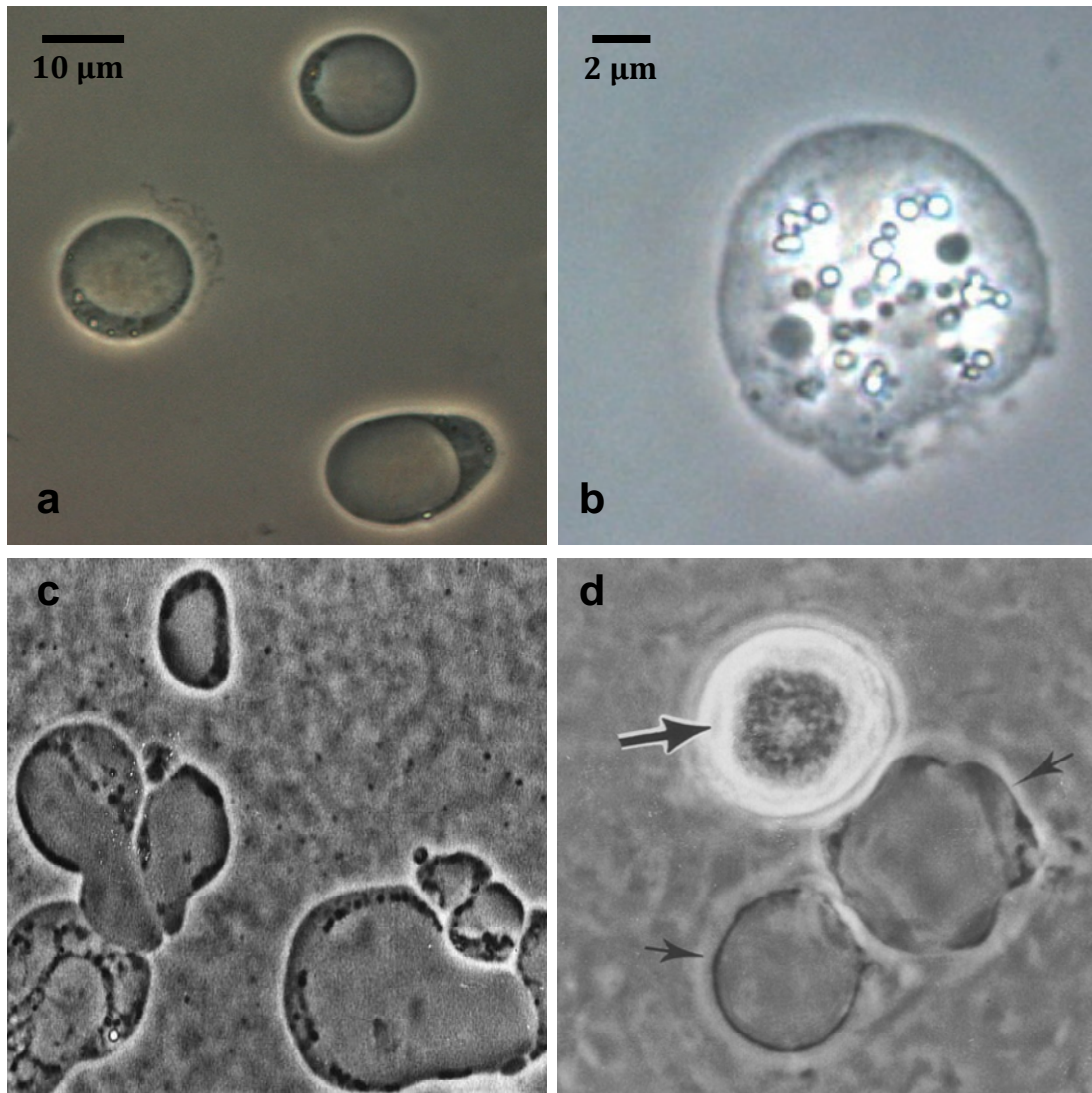


Figure 4 : Observations en microscopie optique en contraste de phase des différentes formes de *Blastocystis* spp. (a) formes vacuolaires en culture axénique. Photo : I. Wawrzyniak (LMGE, UMR CNRS 6023). (b) forme granulaire en culture axénique. Photo : I. Wawrzyniak (LMGE, UMR CNRS 6023). (c) formes amiboïdes de *Blastocystis* spp. obtenues après culture sur milieu IMDM agar. La vacuole centrale est toujours présente. On observe la présence de granules sombres dans le cytoplasme (X400) (Tan SW et al., 1996) (d) : forme kystique de *Blastocystis* (X1000) (grosse flèche noire) et de formes non enkystées (petites flèches noires) (Suresh et al., 1993).

échantillons issus des autres animaux (Parkar et al., 2010). Cette étude fait suite à une étude danoise similaire sur des animaux de zoo et leurs gardiens, montrant une infection des gardiens par le sous-type 8 (25% positifs), sous-type habituellement rare chez l'homme, et retrouvé dans cette étude chez 21% des primates non humains (Stensvold et al., 2009). De même, les isolats de *Blastocystis* spp. humains possèdent la capacité à infecter des animaux de laboratoire (Iguchi et al., 2007; Hussein et al., 2008).

2. *Blastocystis* spp., un parasite polymorphe au cycle de vie incertain

2.1 Un organisme à multiples formes

Blastocystis spp. est un protozoaire polymorphe. Quatre formes majoritaires ont été décrites sur la base d'observations microscopiques issues de culture *in vitro* et d'échantillons de selles : une forme vacuolaire, une forme granulaire, une forme amiboïde et une forme kystique (Figure 4) (Stenzel et Boreham, 1996; Tan, 2008). Plusieurs autres formes, rencontrées moins fréquemment, ont été décrites : une forme avacuolaire, une forme multi-vacuolaire ou encore des cellules présentant des filaments (Stenzel et Boreham, 1996). Il y a cependant très peu d'études, depuis celle de Zierdt (Zierdt, 1991), donnant une description claire de ces formes et vérifiant la présence de celles-ci dans les cultures *in vitro* ou dans les selles. De plus, la multitude des formes présentes rend difficile l'assignation d'une forme spécifique au parasite. Il est en effet difficile de faire la distinction entre une éventuelle nouvelle forme du parasite, et des différences morphologiques et physiologiques induites sur le parasite par un microenvironnement.

L'observation en microscopie électronique à transmission met en évidence des organites typiquement eucaryotes, les plus commodément repérables étant le(s) noyau(x), l'appareil de Golgi, et le réticulum endoplasmique. On constate également la présence d'organites de type mitochondrie, ce qui paraît naturellement surprenant chez

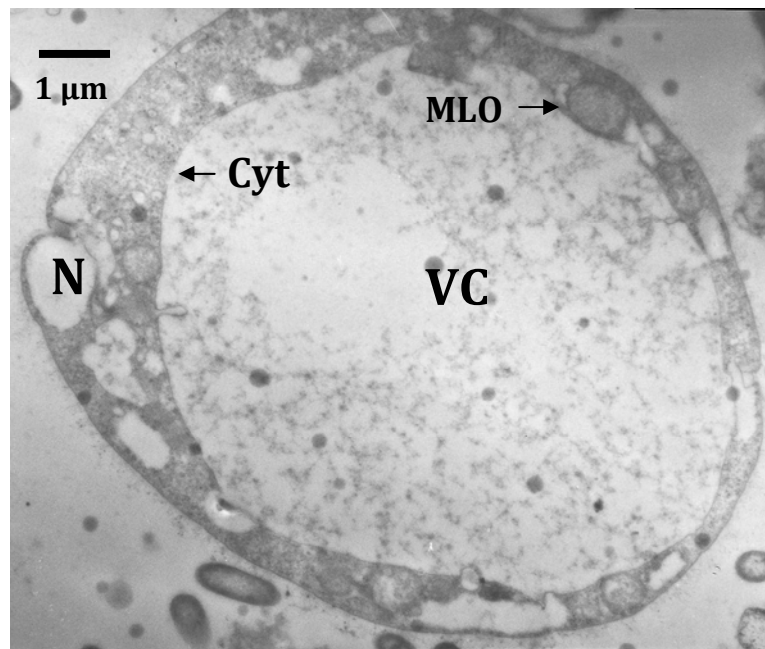


Figure 5: Photographie en microscopie électronique à transmission d'une forme vacuolaire de *Blastocystis* sp. Cette forme est caractérisée par une large vacuole centrale (VC), entourée par une fine bande de cytoplasme (Cyt) où se situent le noyau (N) et les organites apparentés aux mitochondries ou MLO pour « Mitochondrion Like Organelles ». Photo : I. Wawrzyniak (LMGE, UMR CNRS 6023).

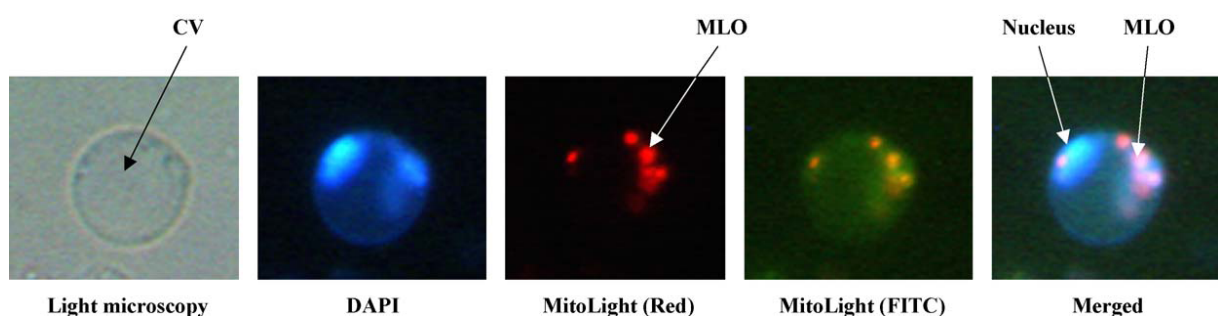


Figure 6: Photographie en microscopie optique d'une forme vacuolaire de *Blastocystis* sp.ST7 en contraste de phase et après coloration au DAPI et au MitoLight. Les colorations au DAPI et au MitoLight mettent en évidence la présence d'ADN dans les organites et un potentiel transmembranaire dans les MLOs (Nasirudeen & Tan 2004).

un organisme anaérobie. Ces organites, décrits pour la première fois par Zierdt (Zierdt, 1986), ont été référencés comme des mitochondries anaérobies. D'un point de vue morphologique ces organites font en moyenne de 0,5 à 1 μm et ont un profil rond légèrement allongé (Figure 5). Ils sont répartis dans le cytoplasme, le plus souvent autour du noyau. Leur nombre peut varier de 1 à 4 dans les parasites. Leur structure est typique des mitochondries, ils sont entourés par une double membrane, leur matrice est dense aux électrons avec la présence de crêtes. Différentes colorations au DAPI et au MitoLight suggèrent que cet organite possède respectivement un génome (Nasirudeen et Tan, 2004) et un potentiel transmembranaire (Figure 6). Cette mitochondrie anaérobie est rapportée comme fonctionnelle (Nakamura et al., 1996), même si aucune activité des enzymes mitochondriales n'a été détectée. La présence de ces organites est restée longtemps déconcertante. En effet, chez les organismes eucaryotes, les mitochondries ont classiquement pour rôle de produire de l'énergie sous forme d'ATP en présence d'oxygène. Cette production d'ATP se fait par la réduction de l'oxygène à travers le processus de phosphorylation oxydative. Le mode de vie anaérobie du parasite a poussé certains chercheurs à suggérer que la mitochondrie de *Blastocystis* spp. était un hydrogénosome (Boreham et Stenzel, 1993). Les hydrogénosomes sont des organites générateurs d'énergie qui diffèrent des mitochondries par l'absence des cytochromes, du cycle de Krebs, et par la présence d'enzymes capables de produire de l'hydrogène. On les retrouve chez divers organismes eucaryotes anaérobies tel que *Trichomonas* et chez certains champignons. Il a également été suggéré que la mitochondrie de *Blastocystis* spp. ne possédait pas de cytochrome (Nakamura et al., 1996).

2.1.1 La forme vacuolaire

La forme vacuolaire est la forme majoritairement observée en culture *in vitro* axénique. De forme sphérique, sa taille peut varier considérablement, allant d'un

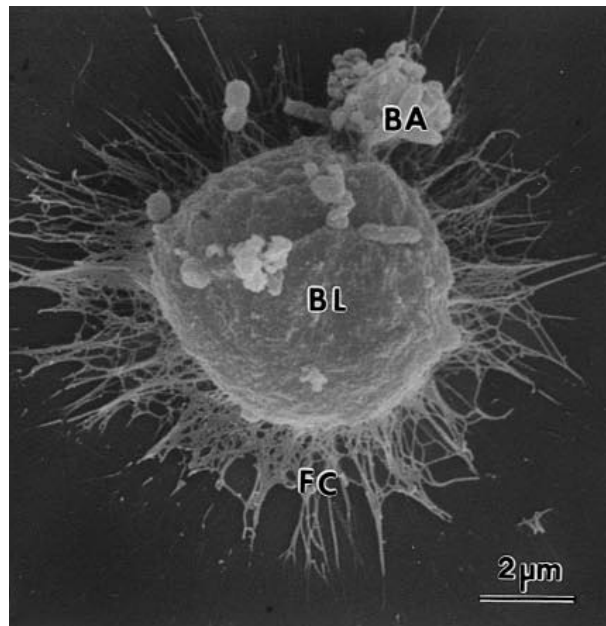


Figure 7 : Photographie en microscopie électronique à balayage de *Blastocystis* sp.

On note la présence de bactéries (BA) sur le glycocalyx (FC) de *Blastocystis* sp. (BL). Le glycocalyx a un aspect fibrillaire (Zaman et al., 1999).

diamètre de 2 μm à 200 μm , pour une moyenne d'environ 15 μm (Figure 4a). Cette forme est caractérisée par une large vacuole centrale pouvant occuper jusqu'à 90% du volume cellulaire. La vacuole repousse le cytoplasme en une fine bande à la périphérie de la cellule. En microscopie électronique à transmission (MET), cette vacuole apparaît claire aux électrons ou peut présenter un aspect finement granuleux distribué de manière irrégulière (Figure 5) (Stenzel et Boreham, 1996; Tan, 2008). La fonction exacte de la vacuole centrale est encore inconnue. Elle pourrait agir comme vacuole de stockage pour stocker des glucides ou des lipides nécessaires à la croissance cellulaire (Yoshikawa et al., 1995a; Yoshikawa et al., 1995b). Un glycocalyx peut être présent à la surface du parasite (Zierdt, 1991). Il posséderait un rôle protecteur contre les chocs osmotiques et le système immunitaire. Il pourrait également intervenir pour faciliter l'adhérence du parasite au mucus de l'épithélium intestinal et être impliqué dans la capture et la dégradation de bactéries à des fins de nutriments (Figure 7) (Zaman et al., 1997; Zaman et al., 1999). Ce dernier phénomène expliquerait l'absence de ce manteau de surface dans les cultures axéniques, qui serait ainsi liée à la présence suffisante de nutriments dans le milieu de culture.

2.1.2 La forme granulaire

La forme granulaire est caractérisée par la présence de granules dans le cytoplasme ou la vacuole centrale (Figure 4b). Sa taille est légèrement plus petite que celle de la forme vacuolaire, et varie de 3 à 80 μm . Les granules ont des aspects hétérogènes et ont été décrits comme de petites vésicules ou des gouttes lipidiques (Dunn et al., 1989). Ces formes sont fréquemment observées dans les cultures non axéniques ou âgées. La modification des conditions de culture peut également faire apparaître ces formes, notamment l'augmentation des concentrations en sérum ou l'addition d'antibiotiques ou d'antifongiques.

2.1.3 La forme amiboïde

La forme amiboïde a été peu décrite, et ses observations restent contradictoires. Celle-ci a été décrite dans des cultures âgées, traitées aux antibiotiques ou dans des échantillons de selles. Elles sont plus petites, jusqu'à 10 µm de diamètre, présentent un aspect irrégulier et possèdent des pseudopodes (Figure 4c) (Dunn et al., 1989; Tan et al., 1996). La vacuole centrale est absente, ainsi que l'appareil de Golgi, le manteau de surface et les MLOs (Dunn et al., 1989). Il reste en revanche un noyau au centre de la cellule. D'autres études décrivent ces formes avec une large vacuole centrale ou plusieurs petites vacuoles dans le cytoplasme (Tan et al., 1996; Tan et Suresh, 2006a), les mêmes organites que la forme vacuolaire ainsi qu'un ou deux pseudopodes (Tan et al., 2001). Malgré la présence de pseudopodes, ces formes semblent non motiles. Certains auteurs suggèrent que ces formes amiboïdes, cultivables sur milieu gélosé, seraient impliquées dans la pathogénie de ce parasite (Tan et al., 2000; Tan et Suresh, 2006a, 2006b).

2.1.4 La forme kystique

La forme kystique (Figure 4d) a été trouvée pour la première fois dans les selles d'un patient en 1988 (Mehlhorn, 1988), et sa première description détaillée date de 1991 (Stenzel et Boreham, 1991). Cette forme est retrouvée dans les fèces stockées plusieurs jours, et très occasionnellement en culture. Le kyste est de petite taille, allant de 3 à 5 µm de diamètre. Une paroi multi-couche sous le manteau de surface, formé à partir des débris de celui-ci, entoure la cellule (Stenzel et Boreham, 1991, 1996). Le cytoplasme apparaît condensé, et la présence de plusieurs petites vacuoles a été observée. Des réserves de glycogènes ainsi que des inclusions lipidiques peuvent être présentes. Ces formes sont capables de survivre un mois à température ambiante et 2 mois à 4°C (Yoshikawa et al., 2004b). Des infections expérimentales chez la souris (Moe et al., 1997) et le rat (Yoshikawa et al., 2004b) montrent qu'il s'agit d'une forme infectante.

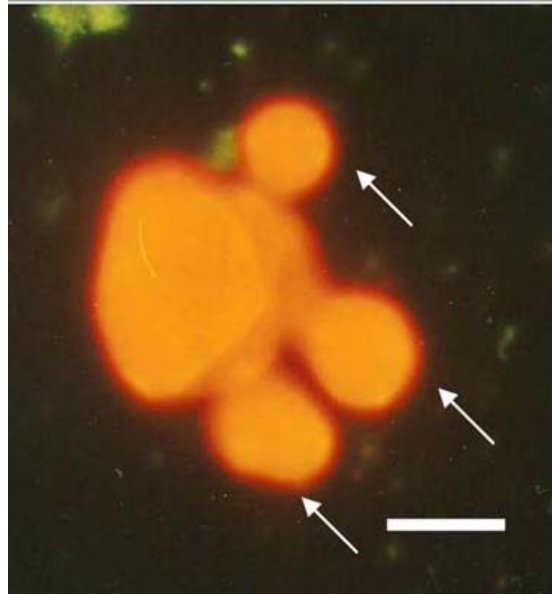


Figure 8 : Photographie d'une multiplication de *Blastocystis* sp. par plasmotomie après coloration à l'acridine orange. On note l'apparition de 3 cellules filles par division du cytoplasme de la cellule mère. Ce phénomène se produit après la division du noyau. Bar = 10µm. (Tan and Suresh, 2007).

2.2 Un cycle de vie encore hypothétique

Le cycle de vie de *Blastocystis* spp. a longtemps été discuté. Plusieurs hypothèses ont été proposées, souvent contradictoires, en raison de la nature pléiomorphe du parasite et de la diversité des modes de reproduction observés : schizogonie⁴ (Singh et al., 1995), plasmotomie (Tan et Suresh, 2007)(Figure 8) ou endodyogénie⁵ (Zhang et al., 2007). La maîtrise de la culture *in vitro* du parasite ainsi que les études menées sur l'observation et la différenciation des formes en culture ont cependant permis l'établissement de cycles hypothétiques. Le cycle de développement actuellement admis est basé sur une division par fission binaire de la forme vacuolaire du parasite, seul mode de division observée en culture *in vitro*. Cependant, en l'absence de vérification *in vivo*, ce cycle de développement reste encore hypothétique (Tan, 2008).

2.2.1 Les différentes formes présentes en culture *in vitro*

Les premiers résultats d'une culture *in vitro* de *Blastocystis* spp. à partir de selles remontent à 1921, dans un milieu xénique constitué d'une solution saline à 0,5% complétement de 10% de sérum humain, à 37°C (Barret, 1921). La croissance du parasite ne s'effectuait qu'au fond des tubes où la concentration en oxygène était la plus faible. Il faudra attendre 53 ans pour obtenir une première culture axénique du parasite, en jarre anaérobie sur un milieu biphasique à base d'œuf, préalablement pré-réduit (Zierdt et Williams, 1974). L'élimination des bactéries a été obtenue après traitement à l'ampicilline, la colistine et la streptomycine. Par la suite un traitement à base de ceftizoxime et vancomycine a été effectué pour éliminer les bactéries résistantes aux premiers antibiotiques (Zierdt, 1991). L'axénisation n'a cependant été un succès que pour certains isolats. Des tentatives s'étant révélées infructueuses, certains chercheurs

⁴ La schizogonie est un mode de reproduction asexuée présent chez certains protozoaires, une cellule multinucléaire se divise de telle sorte que chacune des cellules filles contient un seul des noyaux de la cellule mère.

⁵ L'endodyogénie est un mode de répllication inhabituel où deux cellules filles sont assemblées dans la cellule mère. Une fois les cellules filles complétement formées, la cellule mère se désintègre progressivement.

ont posé l'hypothèse de la nécessité de présence bactérienne pouvant servir de nutriments au parasite ou produisant des métabolites d'intérêts (Boreham et Stenzel, 1993).

Afin de développer un milieu monophasique, de nombreux milieux et conditions ont été testés, tels que le milieu Diamond's TPS-1, le milieu de Jones, le milieu MEM199, le milieu RPMI 1640, le milieu MEM, la supplémentation en sérum bovin ou équin, la réduction des milieux (Stenzel et Boreham, 1996; Tan et al., 2002). Le milieu IMDM supplémenté de 10% de sérum de cheval et pré-réduit 48h s'est révélé être le meilleur milieu monophasique pour cultiver *Blastocystis* spp. (Ho et al., 1993).

De nombreux autres milieux de culture ont également été testés afin d'initier les différentes formes du parasite à partir de la forme vacuolaire. Ainsi la culture du parasite en boîte de Petri sur un milieu IMDM agar semi-solide a permis l'obtention des formes amiboïdes. Ce milieu a également permis l'obtention de souches clonales (Tan et al., 2000). La forme kystique a été obtenue après culture *in vitro* au moyen de différents milieux d'enkystement, le plus efficace étant une solution saline à base de trypticase (Suresh et al., 1993). Il est également possible d'obtenir des formes kystiques en milieu de Jones, à partir d'une culture de parasites issus de selles (Suresh et al., 2009) et de les isoler par centrifugation (Zaman et Khan, 1994).

2.2.2 Cycle de vie

Jusqu'en 1995, le cycle de vie retenu était celui proposé par Zierdt, basé sur plusieurs modes de division : la fission binaire, la schizogonie, la plasmotomie (Zierdt, 1973, 1991). En 1995, Singh et al. ont suggéré un cycle avec un mode de transmission externe, à partir de kystes issus d'une schizogonie d'une forme pré-kystique, et un cycle auto-infectieux conduisant à la multiplication de l'organisme par fission binaire ou schizogonie dans le tractus intestinal de l'hôte. La forme vacuolaire serait à l'origine de ces deux modes d'infections (Singh et al., 1995). En 1996, Stenzel et Boreham (Stenzel et Boreham, 1996), pour qui seule la division binaire est un mode de reproduction

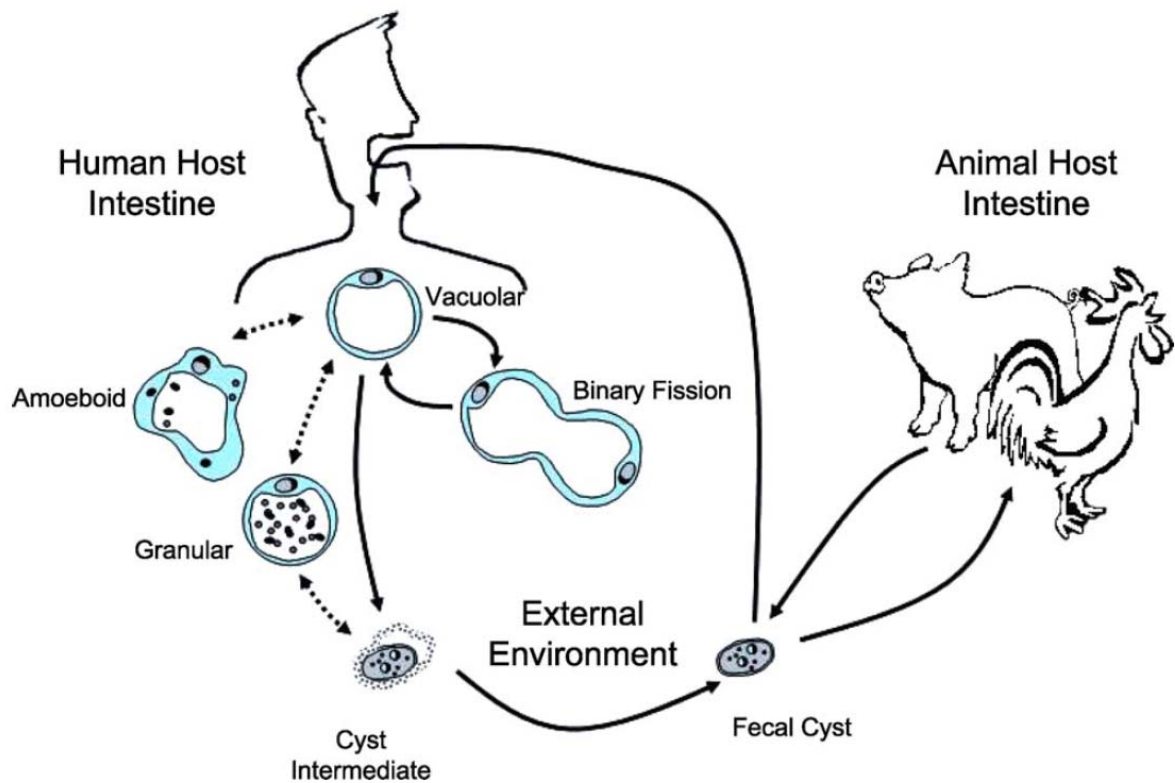


Figure 9 : Cycle de développement de *Blastocystis* spp. proposé par Tan, 2004.

L'infection est initiée par la forme kystique de *Blastocystis* spp. qui est ingérée par un hôte (humain ou animal). Le kyste se différencie alors en formes vacuolaires dans l'intestin, formes qui vont se multiplier par fission binaire. Certaines formes vacuolaires vont ensuite redonner des kystes qui seront disséminés dans l'environnement par l'intermédiaire des fèces, et pourront contaminer un nouvel hôte par transmission oro-fécale. Le développement des formes amiboïdes et granulaires est supposé se faire à partir de la forme vacuaire.

plausible et démontré, propose un nouveau cycle. La forme infestante serait la forme kystique, qui se différencierait en une forme avacuolaire, puis multi-vacuolaire suite à son passage dans l'intestin. La forme vacuolaire serait obtenue uniquement en culture, la vacuole centrale correspondant à une fusion des multiples petites vacuoles. Il faut ensuite attendre 2004, et l'apport de données sur les mécanismes de différenciation de certaines formes pour qu'un nouveau cycle soit proposé (Tan, 2004) (Figure 9). L'infection débiterait par l'ingestion de formes kystiques par l'hôte. Les kystes se développent en formes vacuolaires au niveau du gros intestin et s'y multiplient par fission binaire. Quelques formes vacuolaires s'enkystent et sont éliminées dans les fèces, perpétuant ainsi le cycle. L'apparition et le rôle des formes amiboïdes et granulaires restent cependant à élucider. Ce cycle ne tient pas compte des autres modes de reproduction observés en microscopie, ceux-ci étant jugés artéfactuels par l'auteur. En 2008, Tan (Tan, 2008) modifia son cycle sur la base de données récentes suggérant l'existence de génotypes zoonotiques et en tenant compte des réservoirs possibles pour le parasite (Figure 10). L'homme est infecté par 7 sous-types, dont le sous-type 3 qui semble spécifique à l'homme.

Le cycle de développement de *Blastocystis* spp. mérite donc d'être davantage décrit. Les différents modes de multiplication observés en microscopie tel que l'endodyogénie (Zhang et al., 2007), la plasmotomie (Tan et Suresh, 2007) ou encore la schizogonie (Singh et al., 1995) sont à confirmer pour éventuellement les inclure dans les modèles proposés.

3. *Blastocystis* spp., commensal ou pathogène ?

La pathogénie de *Blastocystis* spp. est controversée, les études menées s'opposant sur son implication ou non dans des désordres gastro-intestinaux. La présence de porteurs asymptomatiques a longtemps fait penser qu'il s'agissait d'un microorganisme

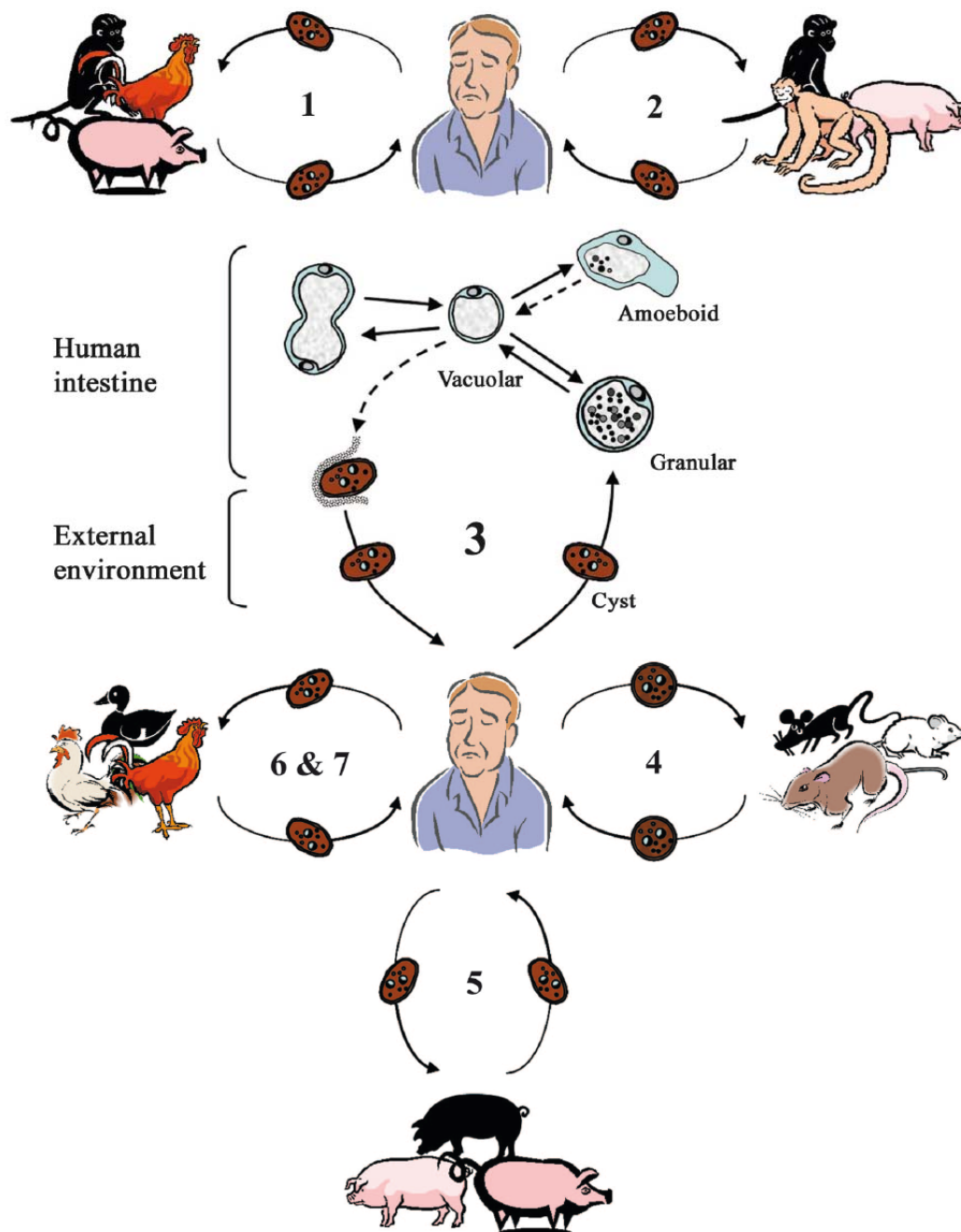


Figure 10 : Cycle de développement de *Blastocystis* spp. modifié par Tan, 2008

Ce cycle est une mise à jour du cycle initialement décrit par Tan (Tan et al, 2004) en incluant les données liées aux sous-types de *Blastocystis* spp., et leur spécificité d'hôte, et prenant également en considération l'aspect zoonotique. L'homme peut être potentiellement infecté par au moins 7 sous-types. Le sous-type 3 (ST3) semble être le seul uniquement retrouvé chez l'homme. Le passage de la forme amiboïde à la forme vacuolaire et de la forme vacuolaire au kyste restent à élucider.

commensal. Pourtant, l'accumulation d'études épidémiologiques, de modèles animaux, de tests *in vitro*, laissent suggérer que *Blastocystis* spp. est pathogène.

3.1 Un organisme associé à des désordres intestinaux chez l'homme

Chez l'homme, les principaux symptômes rapportés sont de types gastro-intestinaux aspécifiques correspondant en majorité à des diarrhées, des douleurs abdominales, des crampes et des nausées. Il a également été rapporté des cas d'anorexie, de perte de poids, de constipation ou de vomissement. Ces symptômes sont d'intensité variable, et les infections peuvent être de type aigu ou chronique (Stenzel et Boreham, 1996; Kaya et al., 2007; Tan, 2008).

Outre ces désordres gastro-intestinaux, il a été rapporté plusieurs cas de manifestations cutanées associées à une infection à *Blastocystis* sp., principalement de l'urticaire de type chronique ou aigu. Cet urticaire a été corrélé avec la présence de *Blastocystis* sp. appartenant aux sous-types 2 et 3 dans les selles des patients (Gupta et Parsi, 2006; Vogelberg et al., 2010; Hameed et al., 2011), ou avec l'observation de forme amiboïde dans le cas du ST3 (Katsarou-Katsari et al., 2008). Il a également été rapporté un cas d'angio-œdème avec une identification de *Blastocystis* sp. dans les selles (Micheloud et al., 2007). Dans tous ces cas les traitements antiparasitaires ont abouti à la résolution de l'infection à *Blastocystis* sp. et à la disparition des lésions cutanées. Les mécanismes par lesquels une infection à *Blastocystis* sp. provoque ces symptômes demandent à être élucidés.

Blastocystis spp. semble également impliqué dans des cas de syndrome de l'intestin irritable (IBS) (Tan, 2008). En effet, une incidence plus élevée des infections à *Blastocystis* spp. chez les patients IBS par rapport à une population contrôle a pu être mise en évidence par certains auteurs (Yakoob et al., 2004; Yakoob et al., 2010b). L'étude d'une population IBS montre une prédominance du ST1 suivi par le ST3. Cette prédominance est inverse à la prédominance dans la population contrôle dans laquelle le sous-type 3 est prépondérant. Les autres sous-types apparaissent identiques dans les

2 populations (Yakoob et al., 2010a). Cependant, si ces études montrent une prévalence plus élevée de *Blastocystis* spp. chez les patients IBS, elles ne permettent pas de conclure que ce parasite puisse être le seul agent étiologique de la maladie. La présence de *Blastocystis* spp. dans les selles de ces patients n'implique pas forcément que les symptômes soient dus uniquement au parasite. Il est envisageable qu'une altération de l'environnement intestinal favorise son développement et qu'il participe de concert avec d'autres facteurs à altérer le fonctionnement de ce tissu. Compte tenu des aspects aspécifiques des symptômes, du facteur psychologique, de la complexité de cet environnement intestinal, la possibilité d'autres causes associées nécessite d'être investiguée. Des études montrant l'éradication de *Blastocystis* spp. concomitante à la disparition des symptômes chez les patients IBS seraient nécessaires pour clarifier l'implication du parasite.

Comme dans beaucoup de parasitoses, les populations immunodéprimées sont plus susceptibles aux infections à *Blastocystis* spp. Il est le protozoaire le plus commun chez les patients VIH ou ayant développé le SIDA, et un des plus fréquents chez les patients ayant déclarés un cancer (Tan et al., 2010).

Dans les cas d'infections symptomatiques à *Blastocystis* spp. et après élimination d'autres facteurs étiologiques, un traitement est habituellement prescrit. Le traitement d'attaque de l'infection est le métronidazole, molécule largement utilisée contre les protozoaires parasites anaérobies (Samuelson, 1999; Gardner et Hill, 2001; Stensvold et al., 2010). Le métronidazole devient actif une fois métabolisé dans les cellules microbiennes. Chez les parasites, le métronidazole est activé quand il reçoit un électron de la ferrédoxine ou de la flavodoxine qui ont eux même été réduits par une pyruvate ferrédoxine oxydo-réductase. Une fois activé, il se lie à l'ADN et provoque une altération de la structure de ce dernier (Samuelson, 1999). La première évaluation de l'efficacité du métronidazole sur *Blastocystis* spp. a été menée en 2003. Cette étude montre que des patients immunocompétents, présentant *Blastocystis* sp. comme seul responsable de diarrhée, répondent parfaitement au traitement (Nigro et al., 2003). Pourtant, il semble

que ce traitement présente une efficacité variable pouvant aller de 0 à 100% (Haresh et al., 1999; Moghaddam et al., 2005; Stensvold et al., 2008; Stensvold et al., 2010). Deux raisons principales aux échecs thérapeutiques ont été proposées : une résistance de certaines souches de *Blastocystis* sp. au métronidazole (Haresh et al., 1999; Mirza et al., 2011), ou une rapide réinfection causée par la non-élimination de la source de contamination, sans que l'on puisse trancher en faveur de l'une ou l'autre. Dans les cas de non-réponse au métronidazole, un traitement à base de trimethoprime/sulfamethoxazole est utilisé comme solution alternative, avec une disparition des parasites dans les selles chez plus de 90% des patients, une disparition des symptômes chez plus de 70% des patients et une diminution des symptômes chez environ 18 % des patients (Ok et al., 1999; Moghaddam et al., 2005; Ertug et al., 2009).

3.2 Des modèles d'études *in vitro* pour évaluer la pathogénie

Les modèles cellulaires constituent aujourd'hui un outil puissant pour étudier les mécanismes impliqués dans une pathologie. Plusieurs études *in vitro* ont ainsi été réalisées afin d'évaluer l'influence du parasite sur les mécanismes cellulaires, notamment sur des cellules épithéliales et pour essayer d'associer *Blastocystis* spp. à des perturbations cellulaires.

Puthia et al (Puthia et al., 2006) ont montré que la co-incubation de cellules épithéliales intestinales de rat IEC6 avec soit des parasites vivants (souche *Blastocystis ratti* WR1) soit des lysats parasitaires, entraîne une entrée en apoptose des cellules. Une modification de la perméabilité cellulaire est également observée, et pourrait être expliquée par le réarrangement ou la rupture des filaments d'actine. *B. ratti* est donc capable de provoquer par contact des modifications cellulaires.

Blastocystis sp. ST4 est également capable de moduler la réponse immune de cellules épithéliales intestinales humaines HT-29 ou T-84 (Long et al., 2001). Ainsi, une culture de 24 h de ces cellules avec le parasite provoque une augmentation de l'excrétion de l'interleukine 8 (IL-8) et une augmentation du facteur de croissance GM-

CSF (granulocyte-macrophage colony stimulating factor). Les IL-8 sont connues pour activer les granulocytes neutrophiles sur un site d'infection mais également les monocytes et les lymphocytes T (Baggiolini et al., 1995), alors que le facteur de croissance GM-CSF intervient dans la stimulation de la production de granulocytes neutrophiles et éosinophiles (Fabian et al., 1992). Les IL-8 et le facteur GM-CSF pourraient jouer un rôle dans le déclenchement de la réponse immune dirigée contre *Blastocystis* sp. ST4, menant de ce fait à une réponse inflammatoire accrue de la muqueuse intestinale. En revanche, aucune augmentation d'IL-8 n'est observée au bout de 6h de co-incubation, laissant supposer la capacité du parasite à diminuer la réponse immunitaire dans les premières phases de l'infection, afin de favoriser cette dernière. De manière indirecte, cet effet pourrait aussi favoriser le développement d'autres agents pathogènes. Cette production d'IL-8 par les cellules T84 semble induite par des protéases à cystéine de manière NF- κ B dépendante, traduisant un rôle inflammatoire de ces protéases (Puthia et al., 2008). Les protéases à cystéine, sont connues pour être présentes et actives chez *Blastocystis* sp. ST7 (Sio et al., 2006), et possèdent une activité variable selon les sous-types (Mirza et Tan, 2009). Elles sont également connues chez les organismes parasites pour jouer un rôle majeur dans la pathogénie (Sajid et McKerrow, 2002). Chez *Blastocystis* sp. ST4 à l'instar d'autres protozoaires comme *Entamoeba histolytica*, les protéases à cystéine ainsi que les protéases à aspartate issues de lysats parasitaires sont capables de dégrader les IgA humaines (Puthia et al., 2005). Ce mécanisme est probablement un des mécanismes permettant la persistance du parasite dans le côlon. Wu et al (Wu et al., 2010) ont caractérisé une protéase à cystéine de la famille des légumaines présente à la surface du parasite. Cette légumaine semble impliquée dans la régulation de l'apoptose du parasite. Ainsi, l'inhibition de son activité entraîne une entrée en apoptose des parasites.

3.3 Des réponses peuvent être apportées par les modèles *in vivo*

Un des obstacles majeurs à l'étude de la pathogénie de *Blastocystis* spp. est l'absence de modèle animal pour vérifier les postulats de Koch. Plusieurs infections expérimentales ont été testées sur différents animaux incluant le rat, la souris ou encore le cochon d'Inde, dans le but de vérifier la capacité d'invasion de la muqueuse intestinale par le parasite et d'associer ce parasite à des symptômes spécifiques notamment l'IBS.

3.4.1 Un modèle limité : la souris

Les différentes infections expérimentales chez la souris ont été faites par voie orale à partir de kystes, ou par voie intra-caecale, intramusculaire et intra-péritonéale à partir de formes vacuolaires en culture (Moe et al., 1997; Moe et al., 1998; Elwakil et Hewedi, 2010). Les effets des infections par voie orale ou intra-caecale restent assez limités : perte de poids, léthargie, difficulté de déplacement (Moe et al., 1997; Elwakil et Hewedi, 2010). Selon l'âge des souris, les infections disparaissent : au bout de 3 semaines pour des souris âgées de 2 semaines au moment de l'infection, et au bout de 2 semaines pour des souris âgées de 4 à 6 semaines au moment de l'infection (Moe et al., 1997). Les souris âgées de 8 semaines se révèlent même toutes insensibles à l'infection (Moe et al., 1997). Au niveau intestinal, des formes vacuolaires et granulaires de *Blastocystis* spp. sont retrouvées dans le gros intestin, principalement au niveau du caecum. La présence de kystes a également été décelée au niveau du côlon (Moe et al., 1997; Elwakil et Hewedi, 2010). Les examens histo-pathologiques réalisés au niveau du caecum ne montrent aucune lésion particulière (Moe et al., 1997). Au niveau de la localisation, Moe et al montrent les parasites présents dans la lumière, et quelques uns à la surface de l'épithélium intestinal (Moe et al., 1997). Toutefois, une pénétration de *Blastocystis* spp. au niveau de la lamina propria et de la sous-muqueuse a aussi été observée, cette pénétration étant accompagnée d'une intense réaction inflammatoire (Elwakil et Hewedi, 2010).

Les expérimentations par injection musculaire de *Blastocystis* spp. ont montré des effets toxiques et nécrosants sur les cellules, accompagnés d'une sévère réaction inflammatoire (Moe et al., 1998). Ces effets restent réversibles, les cellules musculaires se régénérant au bout d'une semaine.

Les expérimentations par injections intra-péritonéale réalisées à partir de parasites issus de selles de patients et cultivés sur RPMI-1640 montrent une pathogénie du parasite dépendante de la dose (Zhou et al., 2010). La mortalité des souris est croissante en fonction du nombre de parasites injectés. Les auteurs observent ainsi une mortalité de 20% chez les souris infectées avec 10^4 parasites, de 30% chez les souris infectées avec 10^5 parasites et de 45% chez celles infectées avec 10^6 parasites (Zhou et al., 2010). Les souris survivantes présentent un comportement apathique, léthargique et un pelage hérissé, les symptômes disparaissant au bout d'une semaine. Trois formes du parasite (vacuolaire, granulaire et kystique) ont été retrouvées dans l'abdomen des souris mortes. Chez les souris vivantes, seuls des kystes ont été retrouvés, laissant supposer un lien entre les formes du parasite et la pathogénie.

La souris apparaît donc comme un modèle expérimental limité puisqu'il semble en effet difficile de reproduire ce qui se passe chez l'homme. Ceci peut être dû à l'utilisation de souches potentiellement peu ou pas pathogènes chez le rongeur ou à une susceptibilité moindre de la souris à *Blastocystis* spp.

3.4.2 Un modèle plus sensible : le rat

Contrairement à la souris, les différentes études menées sur le rat en font un potentiel modèle animal (Tan, 2008). Une centaine de kystes, isolés à partir d'un rat de Norvège (*Rattus norvegicus*) et d'un cochon d'Inde, inoculée par voie orale suffisent à entraîner une infection chez le rat wistar (Yoshikawa et al., 2004b) confirmant une transmission oro-fécale. Comme chez la souris, le caecum et le gros intestin sont infestés alors que le petit intestin semble sain. Plus récemment, des infections expérimentales ont été réalisées à partir de parasites issus de patients et injectés intra-caecalement aux

rats (Iguchi et al., 2007). Cette étude montre une variation du degré des infections, qui est fonction des sous-types testés (2, 3, 4 et 7). Seul le ST3 s'est avéré incapable d'infecter les rats. Des infections expérimentales ont également été réalisées à partir de STs 1, 3 et 4 de *Blastocystis* sp. issus de patients symptomatiques et asymptomatiques et d'un ST2 issu uniquement de patients asymptomatiques (Hussein et al., 2008). Les rats infestés à partir des isolats de patients symptomatiques développent des pathologies modérées à sévères au contraire des rats infestés avec des isolats issus de patients asymptomatiques qui développent peu voire aucune pathologie. Parmi les isolats symptomatiques, seul le ST1 induit une mortalité (25%). Les auteurs concluent que seul le ST1 est un sous-type pathogène, et que le potentiel pathogène des STs 3 et 4 est variable.

Le potentiel pathogène variable entre isolats d'un même sous-type est illustré par une étude sur un sous-type 4 au travers de l'analyse de l'expression d'interféron gamma et des cytokines pro-inflammatoires par la muqueuse caecale de rats infectés expérimentalement (Iguchi et al., 2009). Ainsi malgré une augmentation des transcrits de l'IFN-gamma, de l'interleukine 12 ou des TNF alpha, aucune lésion pathologique n'est apparue sur la surface de la muqueuse, 2 à 3 semaines après l'infection. De même, l'expression des protéines des jonctions serrées suggère que l'intégrité structurelle de l'épithélium caecal n'est pas sévèrement affectée. Il semblerait donc que l'isolat de sous-type 4 étudié dans cette étude ne soit que modérément pathogène.

Le rat semble un modèle animal plus intéressant que la souris. De plus, les observations montrent une sensibilité variable des rats aux différents isolats ou sous-types. Ceci est un facteur intéressant pour les infections expérimentales et nécessiterait des études pour vérifier s'il s'agit d'une sensibilité moindre des rats ou un pouvoir pathogène différent de certains sous-types de *Blastocystis* sp.

3.5 *Blastocystis* spp., un parasite pathogène ?

Blastocystis spp. apparaît aujourd'hui de plus en plus comme un pathogène émergent. En effet, l'apparition de nouvelles méthodes de diagnostic a mis en évidence une forte prévalence de *Blastocystis* spp. dans les populations humaines. L'existence d'isolats résistants aux traitements thérapeutiques, son mode de contamination (oro-fécale) et son potentiel zoonotique facilitent grandement sa dissémination. De plus, *Blastocystis* spp. dispose de quelques atouts lui permettant non seulement de survivre dans l'environnement intestinal, mais également d'y persister. Ainsi, il (1) dispose d'un métabolisme anaérobie adapté aux tensions faibles en oxygène dans l'environnement intestinal ; (2) possède des protéases capables de désorganiser les cellules intestinales; (3) est capable de dégrader des IgA et de moduler la réponse immunitaire, le protégeant ainsi des effecteurs de l'immunité. De plus, il est montré associé à de nombreux symptômes gastro-intestinaux, et des lésions cutanées chez des patients symptomatiques. Toutes ces données renforcent l'idée que *Blastocystis* spp. a certainement plus d'importance clinique qu'estimé actuellement. Dès lors, il semblait intéressant d'essayer de clarifier son pouvoir pathogène et d'identifier les facteurs potentiels impliqués dans sa pathogénie. Pour cela, nous nous sommes lancés dans le décryptage du génome d'un sous-type 7 de *Blastocystis* sp. Ces données ont été utilisées pour identifier *in silico* des gènes candidats pouvant être impliqués dans les processus infectieux et pathogénique de ce parasite. Ce travail a ouvert la voie à la caractérisation de protéases sécrétées par le parasite et pouvant jouer un rôle majeur dans la physiopathologie de la blastocystose. Ce travail a également permis de fournir des données sur le métabolisme des MLOs.

II. La génomique, une avancée dans la connaissance des organismes

La génomique a révolutionné l'étude des organismes et a ouvert de formidables perspectives sur le plan scientifique. Cette discipline étudie les génomes tant au niveau du nombre, de la disposition ou de la séquence des gènes qu'au niveau des produits et de la fonction de ceux-ci. Le déchiffrement de ces génomes complets associé aux données de transcriptomique, de protéomique ou de métabolomique a engendré une profusion de données complexes qui ont modifié les approches liées à la connaissance des organismes et largement aidé à la compréhension de leur biologie. Ainsi, ces données, qui offrent un vaste champ d'investigation aux questions de biologie fondamentale, permettent d'étudier les entités dans leur ensemble (réseaux géniques, interactions protéiques...) mais offrent également la possibilité d'éclaircir les principes sous-jacents au développement, à l'adaptation aux conditions environnementales, ou encore à la filiation évolutive des organismes.

1. La génomique pour l'étude des parasites.

Bien que les génomes eucaryotes soient souvent complexes (grand nombre de gènes composés d'introns, existence de longues régions inter-géniques, présence de séquences répétées) et de grande taille, ces dernières années ont vu leur séquençage se multiplier. Ceci a été rendu possible par les avancées technologiques ayant permis la réduction du coût et du temps nécessaire au séquençage, à l'assemblage et à l'annotation. Parmi les multiples génomes séquencés ou en cours de séquençage, beaucoup correspondent à des protozoaires ou champignons parasites d'intérêts médical ou vétérinaire, dont certains sont la cause de forte mortalité chez l'homme. On peut ainsi citer la microsporidie *Encephalitozoon cuniculi* premier parasite eucaryote séquencé (Katinka et al., 2001), *Trypanosoma brucei* (Berriman et al., 2005) responsable

Tableau 4 : Liste des principaux génomes de protozoaires parasites de l'homme séquencés.

Organisme	Taille du génome en Mb	Nombre de gène prédits	Réf
<i>Cryptosporidium hominis</i>	10,4	3994	Xu et al, 2004
<i>Cryptosporidium parvum</i>	16,5	3807	Abrahamsen et al, 2004
<i>Plasmodium falciparum</i>	22,9	5268	Gardner et al, 2002
<i>Plasmodium vivax</i>	26,8	5433	Carlton et al, 2008
<i>Giardia Lamblia</i>	11,7	6470	Morrison et al, 2007
<i>Leishmania braziliensis</i>	33	8314	Peacock et al, 2007
<i>Leishmania infantum</i>	33	8195	Peacock et al, 2007
<i>Leishmania major</i>	32,8	8272	Ivens et al, 2005
<i>Trichomonas vaginalis</i>	160	59681	Carlton et al, 2007
<i>Trypanosoma brucei</i>	26	9068	Berriman et al, 2005
<i>Trypanosoma cruzi</i>	34	22570	El-Sayed et al, 2005
<i>Entamoeba histolytica</i>	23,8	9938	Loftus et al, 2005
<i>Giardia intestinalis</i>	11,7	4691	Franzen et al, 2009

de la maladie du sommeil, *Plasmodium falciparum* responsable du paludisme (Gardner et al., 2002) ou encore *Entamoeba histolytica* agent de la dysenterie amibienne (Loftus et al., 2005) (Tableau 4). Le séquençage de ces génomes, ainsi que l'identification et la caractérisation des gènes les constituant, ont considérablement enrichi les connaissances sur l'organisation des réseaux géniques, les interactions protéiques, les voies physiologiques mais également sur l'évolution de ces génomes. Ces données s'avèrent offrir d'immenses possibilités pour la recherche sur ces organismes. Pour autant quel a été l'impact de la génomique sur la parasitologie ?

1.1 L'intérêt de séquencer un génome complet

Disposer des données issues du séquençage de génomes complets a joué un rôle essentiel dans la compréhension des potentialités des parasites. De nombreuses études témoignent du rôle prépondérant de ces données dans la détermination de gènes d'intérêts et de voies géniques qui peuvent représenter des cibles thérapeutiques potentielles, ou de candidats pour l'élaboration d'un vaccin ou d'outils de diagnostic (Carlton, 2003). Par exemple, l'analyse du génome de *Plasmodium falciparum* et la recherche d'homologies à partir de gènes à fonction connue chez des plantes a permis d'identifier plusieurs gènes impliqués dans les voies de synthèse de type II des acides gras (Waller et al., 1998; Waller et al., 2003). Cette voie métabolique localisée au niveau de l'apicoplaste (une relique du chloroplaste) est commune aux chloroplastes des plantes et aux bactéries mais distincte de la voie de synthèse de type I des acides gras chez les mammifères y compris l'homme. Des inhibiteurs spécifiques de cette voie métabolique comme le triclosan, le thiolactomycine ou leurs analogues ont été démontrés comme ciblant cette voie métabolique et inhibant le parasite, offrant ainsi de nouvelles pistes pour le développement de drogues anti-malaria (Surolia et Surolia, 2001).

Les données de séquençage permettent aussi l'identification de gènes ou de familles multi-géniques codant des protéines essentielles au cycle de développement des

parasites, impliquées dans la mise en place de structures cellulaires particulières ou dans les mécanismes de virulence. Chez le parasite intestinal *Giardia lamblia*, la majorité des gènes codant les protéines de la paroi kystique a été identifiée à partir des séquences du génome complet (Sun et al., 2003). En effet, les infections à *Giardia* débutant par l'ingestion de kystes, la mise en place de la paroi kystique est essentielle pour la survie du parasite lors de son transit dans l'estomac. L'identification de ces protéines peut conduire à l'élaboration de drogues capables d'interférer lors de l'assemblage de la paroi kystique et ainsi bloquer le cycle de vie du parasite.

La disponibilité, en plus des génomes de parasites, des génomes des hôtes ou des vecteurs, fournit une richesse supplémentaire pour l'analyse et la compréhension des interactions vecteurs/parasites et hôte/parasites. Les efforts menés pour intégrer toutes ces données de séquences ont amené à une meilleure compréhension des interactions entre ces organismes, au niveau du déroulement de la transmission du parasite, de la susceptibilité de l'hôte, de l'immunité de l'hôte et du vecteur, de la pathogénie du parasite (Carlton, 2003). Cette intégration peut être vue comme la constitution d'un système biologique visant à expliquer le rapport complexe entre un génotype et un phénotype en employant ce réseau d'interaction. Cette approche a notamment été un succès dans la reconstruction de réseau métabolique.

1.2 La reconstruction métabolique

L'annotation d'un génome et la reconstruction métabolique sont devenus complémentaires. La reconstruction métabolique va apporter un sens biologique aux informations issues de l'annotation d'un génome en intégrant l'ensemble des protéines de ce génome dans des réseaux métaboliques. Cela va permettre d'établir un pont entre ces informations issues du séquençage, les analyses de ce génome, et les multiples données biochimiques disponibles, mais également permettre de compléter et corriger les données génomiques. Cette démarche très complexe de reconstruction métabolique nécessite l'utilisation de bases de données aussi bien généralistes comme NCBI, UNIPROT,

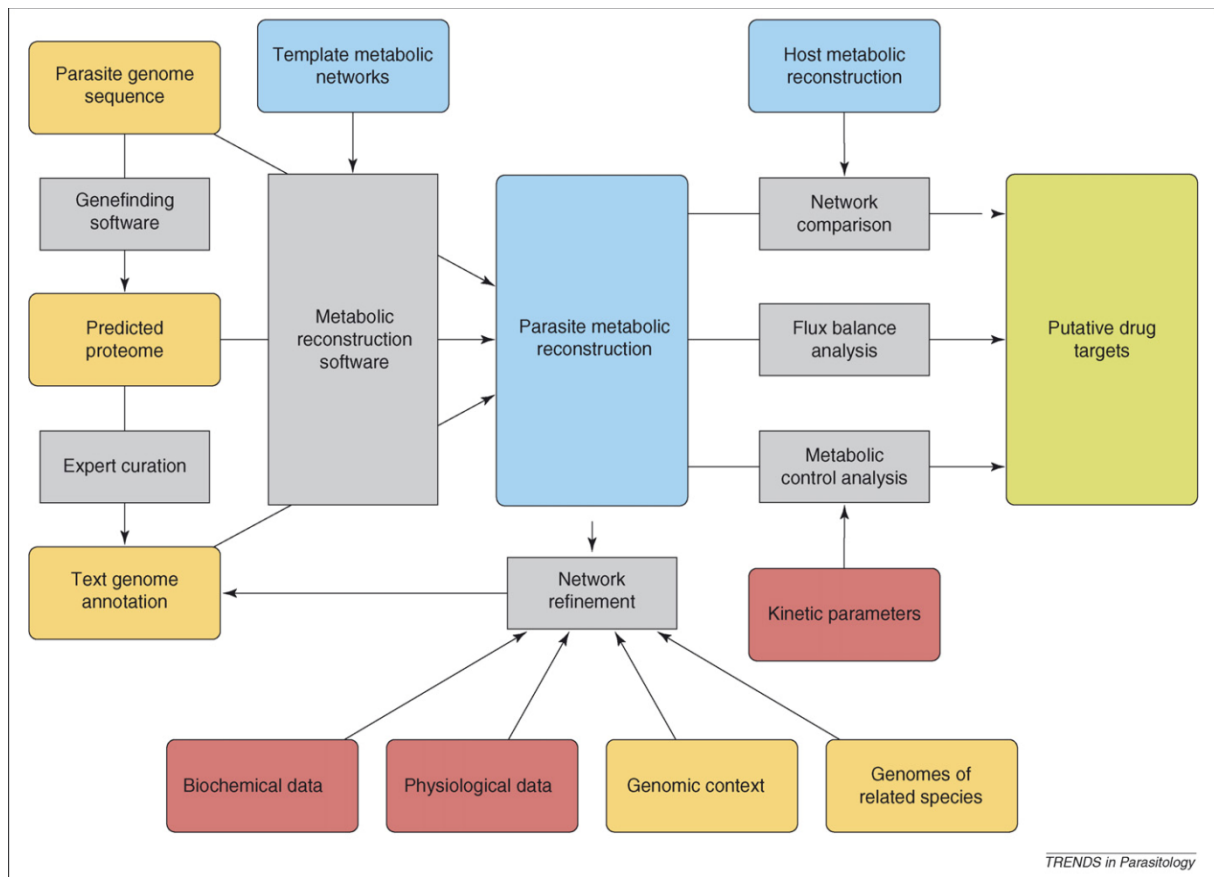


Figure 11 : Schéma illustrant les principales données et outils permettant la reconstruction métabolique. L'organigramme montre les données (rectangles arrondis) et les opérations (rectangles gris) nécessaires à la reconstruction. Les couleurs représentent la nature des données : en rouge les données expérimentales ; en orange les données dérivées des séquences du génome ; en bleu les reconstructions de réseaux ou de voies métaboliques ; en vert les cibles potentielles de drogues (Pinney et al., 2007). La reconstruction métabolique apparaît comme le point de convergence entre les informations issus des analyses *in silico* des génomes et celles apportées par les procédés expérimentaux tels que la transcriptomique ou la protéomique. On obtient ainsi une complémentarité entre les différents types d'informations.

que spécialisées comme KEGG, mais peut également intégrer des données expérimentales biochimiques ou physiologiques issues d'analyses protéomiques ou transcriptomiques (Figure 11) (Pinney et al., 2007). La reconstruction métabolique va ainsi éclairer sur les possibilités métaboliques d'un organisme, notamment concernant son adaptation à son environnement ou à son hôte, données d'autant plus importantes pour les parasites puisqu'ils tirent avantage du métabolisme de l'hôte. Ces données permettent d'apporter des corrections à l'annotation des génomes, ainsi la présence d'une fonction enzymatique isolée peut indiquer une mauvaise annotation de gènes, de même qu'une enzyme manquante au sein d'une voie métabolique semblant complète peut indiquer la présence d'un gène dont la fonction n'a pas été assignée mais dont l'existence s'impose pour le bon fonctionnement du modèle. Ces données offrent aussi la possibilité d'effectuer la comparaison de métabolismes intra et inter-espèces (Durot et al., 2009).

Une application directe de la reconstruction métabolique chez les parasites à partir des données génomiques est la recherche de nouvelles cibles de drogues (Apic et al., 2005). En effet, connaître les voies métaboliques d'un organisme peut servir de cadre pour trouver des cibles potentielles et évaluer l'action d'une drogue sur ces cibles. Les voies métaboliques situent la cible dans son contexte, donnent des indications sur ses interconnexions dans la cellule en indiquant les protéines en amont et en aval dans les voies de signalisation. Ceci permet de déterminer *in silico* les réactions en cas d'interférence de la cible : les effets descendants bénéficiaires ou non, l'existence de voies métaboliques pouvant compenser les interférences.... Toutes ces connaissances peuvent donner des indications pour aider au choix de la cible.

A l'échelle d'un système biologique (intégration des données génomiques de l'hôte et du parasite), la comparaison des métabolismes de l'hôte et du pathogène apportent des informations encore plus précises : au niveau des interactions entre l'hôte et le parasite, de l'évolution des organismes, mais également d'un point de vue des recherches de cibles thérapeutiques (Pinney et al., 2007). Les différences de gènes

observées entre une même voie métabolique chez l'hôte et le parasite peuvent être dues à la perte ou au gain de gènes. Beaucoup de parasites ont ainsi perdu des métabolismes nécessaires aux organismes libres, prélevant les nutriments nécessaires chez leurs hôtes. *Encephalitozoon cuniculi* par exemple, dont le génome est extrêmement compact, est fortement dépendant de son hôte, en raison notamment de l'absence de gènes au niveau de certaines voies métaboliques dont le cycle de Krebs (Katinka et al., 2001). Dans le même ordre d'idée, certains parasites possèdent des fonctions métaboliques absentes chez l'homme (Chaudhary et Roos, 2005). On peut citer en exemple la dégradation de l'hémoglobine par les plasmodies qui fait appel à des protéases spécialisées (falcipaine) (Fairlamb et Cerami, 1992). Ces fonctions métaboliques spécifiques aux parasites sont des cibles très attractives pour la recherche de nouvelles drogues.

1.3 La génomique comparative

L'analyse comparative de plusieurs génomes a permis d'apporter de nouvelles données intéressantes pour comprendre les différences biologiques. Elle couvre un vaste champ de recherche qui compare la composition globale d'un génome, l'organisation chromosomique, la synténie de régions conservées, l'identification de gènes orthologues et paralogues, l'étude de gènes spécifiques d'une espèce (Carlton, 2003). Les projets de séquençage tendent maintenant vers le séquençage simultané de plusieurs organismes ou espèces phylogénétiquement proches. Le séquençage complet des différentes espèces d'*Entamoeba* : *Entamoeba dispar*, *E. invadens* (Wang et al., 2003) et *E. histolytica* (Loftus et al., 2005) en est une illustration.

Pourtant, la génomique comparative reste une discipline récente chez les eucaryotes, la première comparaison de deux génomes complets remonte à 2002 avec la comparaison de deux pathogènes du même genre, *Plasmodium falciparum* responsable de la malaria humaine et *Plasmodium yoelii yoelii* responsable de la même maladie chez le rongeur (Carlton et al., 2002). La comparaison de la composition globale des deux

génomomes a montré plusieurs similarités. Les deux génomes sont haploïdes, composés de 14 chromosomes linéaires, avec 60% de gènes orthologues entre les deux espèces. La comparaison des gènes prédits révèle que les deux espèces possèdent des caractéristiques similaires dans les régions codantes et non codantes, comme le pourcentage en adénine et thymine (% A/T). Cette comparaison a également permis de mettre en évidence les gènes spécifiques de chaque espèce, notamment au niveau des familles de gènes impliquées dans la variation antigénique ou l'évasion du système immunitaire. Ces mécanismes de défense développés par ces deux espèces du genre *Plasmodium* ont ainsi évolué en fonction de l'environnement hostile et des réponses actives du système immunitaire de l'hôte. Pour compléter l'analyse comparative des deux espèces, une carte de la synténie des génomes a été réalisée, montrant des régions de plusieurs kilobases conservées à la fois dans l'ordre, mais aussi dans la structure des gènes. Les données générées par l'étude de la synténie peuvent apporter des informations pour la prédiction et l'annotation de gènes. Les régions synténiques dans le génome de deux espèces peuvent également être utilisées comme méthode pour déterminer des gènes orthologues (Carlton, 2003).

La génomique comparative apporte également des informations du point de vue de l'évolution des organismes. Une base fondamentale de la génomique comparative est que les génomes comparés partagent un ancêtre commun, les différences observées entre ces génomes peuvent donc être expliquées par les pressions de sélections (Carlton, 2003). En reprenant l'exemple des plasmodies, la comparaison des génomes de *P. falciparum* et des espèces responsables de la malaria chez les rongeurs (*P. Yoelii*, *P. berghei*, *P. chabadi*) a mis en évidence un ensemble de 4500 orthologues situés dans les régions centrales fortement synténiques des chromosomes, et qui ont permis de délimiter les régions sub-télomériques variables (Kooij et al., 2005). La génération d'une carte de synténie entre ces génomes a permis de définir que les régions centrales des 14 chromosomes de *P. falciparum* et des espèces des rongeurs étaient organisées en 36 blocs synténiques. Ces blocs représentent des groupes de gènes qui ont été hérités

depuis que ces espèces ont divergé, mais dont l'organisation a changé en raison d'événements de recombinaisons (Kooij et al., 2005).

La génomique comparative permet donc d'enrichir les connaissances sur les espèces, aussi bien en terme d'acquisition de diversité due aux pressions de sélection exercées par l'hôte, un vecteur ou l'environnement, qu'en terme de composition globale des génomes.

2. Les Straménopiles, classification et génomique

Comme nous l'avons vu précédemment, séquencer un génome peut apporter une masse importante de données. Dans le cas de *Blastocystis* spp. plusieurs génomes et des banques d'EST de Straménopiles étaient disponibles, nous offrant ainsi des données utiles à l'annotation d'un nouveau génome de Straménopiles.

2.1 Classification d'un phylum très hétérogène

Le phylum des Straménopiles fut établi par Cavalier-Smith en 1986 (Cavalier-Smith, 1986), regroupant des eucaryotes motiles bi-flagellés avec un flagelle à mastigonème, et tous leurs descendants ayant perdu un ou l'ensemble de leurs flagelles. Ils sont présents dans les eaux douces, dans le milieu marin, ou dans les habitats terrestres. La diversité au sein des Straménopiles est frappante, on y retrouve de grandes algues pluricellulaires jusqu'à de petits organismes unicellulaires, des organismes photosynthétiques et non photosynthétiques, des libres et des parasites. En raison de cette grande diversité, les Straménopiles ont été scindés en trois sous-phylums : les Ochrophyta regroupant les autotrophes, les Bigyra parmi lequel se retrouve *Blastocystis* spp. et les Pseudofungi parmi lesquels se retrouvent de nombreux organismes non photosynthétiques (Cavalier-Smith et Chao, 2006). Une étude phylogénique de 35 taxa appartenant aux Ochrophyta et aux hétérotrophes à partir des séquences des ARNr des petites et grandes sous-unités ribosomiques montre les

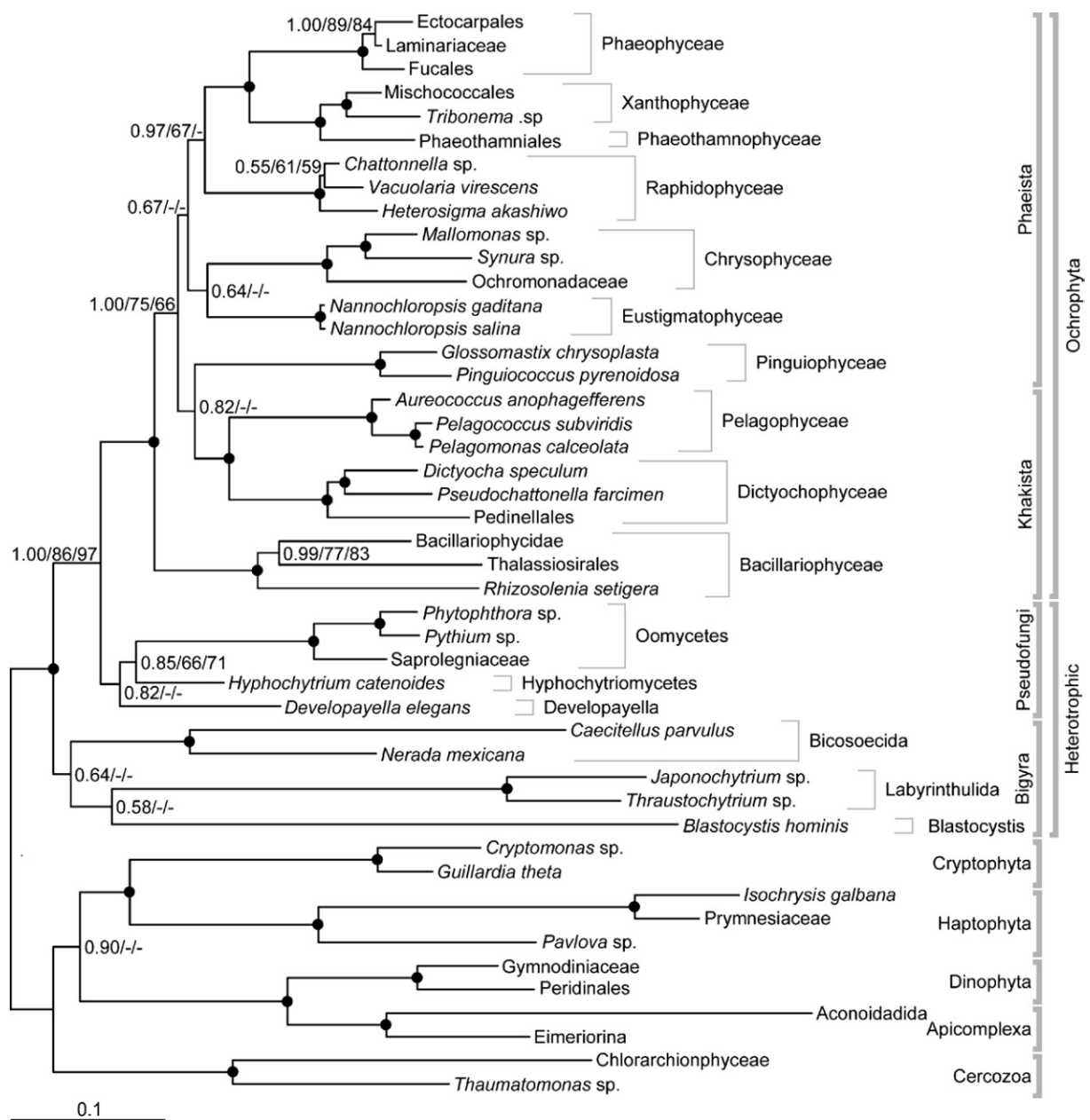


Figure 12 : Arbre phylogénétique des Straménopiles basé sur les séquences ARNr 18S et 28S. L'arbre est construit selon la méthode des inférences de Bayesian. Les valeurs des nœuds internes correspondent à la probabilité de l'événement (inférence de Bayesian), et aux valeurs de bootstrap pour Treefinder et RAXML. Les ronds noirs indiquent une probabilité de 1 et un bootstrap supérieur à 90% (Riisberg et al., 2009).

premiers comme un sous-phylum monophylétique distinct de celui des hétérotrophes (Riisberg et al., 2009). En revanche les deux embranchements des hétérotrophes restent faiblement appuyés par les valeurs de bootstrap (Figure 12), laissant supposer que l'ajout de séquences supplémentaires pourraient modifier l'organisation phylogénétique des deux clades des hétérotrophes.

2.2 La génomique des Straménopiles

Certains Straménopiles ont fait l'objet de projets de séquençage de leurs génomes (Tableau 5). Les premiers projets ont concerné plusieurs espèces de *Phytophthora* spp. en raison de l'importance des dégâts causés par ces parasites sur des plantes d'intérêt agronomique comme le soja ou la pomme de terre. Le séquençage de génomes complets de *P. sojae* et *P. ramorum* a été entrepris par le Joint Genome Institute (JGI, Etats-Unis) en 2004, par la méthode de shotgun total. Le génome de *P. sojae* d'une taille de 95 Mb a été séquencé avec une couverture de 9x, et est composé de 19027 gènes prédits. Celui de *P. ramorum* possède lui une taille de 66 Mb, pour une couverture de 7x, et est composé de 15743 gènes (Tyler et al., 2006). La disponibilité de plusieurs banques d'EST de différentes espèces de *Phytophthora* spp. a largement facilité l'annotation de ces génomes (Qutob et al., 2000; Torto et al., 2003; Randall et al., 2005). Par la suite le séquençage du génome de *P. infestans* a été entrepris au Broad Institute (Etats-Unis). Ce génome d'environ 240 Mb comporte 17797 gènes prédits, et a été couvert 8x (Haas et al., 2009). Il s'agit du plus grand et du génome le plus complexe séquencé parmi les Achromolveolata. Ce génome présente de multiples séquences répétées, ayant rendu son assemblage difficile. La comparaison avec les 2 autres génomes de *Phytophthora* montre un nombre plus important de gènes codant des protéines sécrétées effectrices de la maladie et prédites pour avoir une activité altérant la physiologie de l'hôte, ainsi que plusieurs gènes induits durant l'infection (Haas et al., 2009). La comparaison des trois génomes révèle également une organisation peu commune, comprenant des blocs de gènes conservés y compris dans leur ordre. Ces blocs présentent une forte densité

Tableau 5 : Liste des génomes de Straménopiles séquencés

Organisme	statut	Taille du génom en Mb	Nombre de gène prédicts	Méthode de séquençage	couverture	Ref
<i>Phytophthora infestans</i>	organisme parasite de plante (oomycète)	240	17797	whole genome shotgun	8X	Haas et al, 2009
<i>Phytophthora sojae</i>	organisme parasite de plante (oomycète)	95	19027	whole genome shotgun	9X	Tyler et al, 2006
<i>Phytophthora ramorum</i>	organisme parasite de plante (oomycète)	66	15743	whole genome shotgun	7X	Tyler et al, 2006
<i>Thalassiosira pseudonana</i>	organisme libre (diatomée)	34,5	11242	whole genome shotgun	14X	Amburst et al, 2004
<i>Phaedodactylum tricornutum</i>	organisme libre (diatomée)	27,4	10402	whole genome shotgun	18X	Bowler et al, 2008

génique et peu de répétitions. Ils sont séparés par des régions dans lesquels l'ordre des gènes n'est pas conservé, la densité génique est faible et les répétitions élevées. Ces blocs regroupent 90% des orthologues des 3 espèces, représentant 70% des gènes de *P. infestans* et 78% des gènes de *P. sojae* et *P. ramorum*. Les régions entre les blocs conservés présentent des distances intergéniques plus grandes et augmentant selon la taille du génome (1,5 kb pour *P. ramorum*, 2,2 kb pour *P. sojae*, 3,7 kb pour *P. infestans*). Les gènes impliqués dans la pathogénie se situent principalement dans ces régions. Ceci a probablement joué un rôle crucial dans l'adaptabilité du pathogène à la plante et dans son potentiel évolutif (Haas et al., 2009).

En plus des génomes de Straménopiles parasites que sont les *Phytophthora*, deux génomes de diatomées (organismes importants en écologie) ont également été séquencés, *Thalassiosira pseudonana* (Armbrust et al., 2004) et *Phaeodactylum tricornutum* (Bowler et al., 2008). La diatomée *Thalassiosira pseudonana*, premier eucaryote phytoplanctonique marin séquencé, fut choisie en raison de son appartenance à un genre distribué dans les océans du monde entier et parce que cette espèce a servi de modèle d'étude de la physiologie des diatomées. Son génome fut séquencé par shotgun total, avec une couverture de 14x, pour une taille de 34,5 Mb et 11242 gènes prédits. Plus de la moitié des gènes n'ont pas de fonction assignée. L'analyse du génome a cependant permis de reconstruire en partie le métabolisme de cette diatomée adaptée à la vie en milieu marin. Résultat surprenant, *T. pseudonana* présente un cycle de l'urée et un métabolisme des acides gras dans les mitochondries, deux caractéristiques rencontrées jusqu'ici seulement chez les animaux, jamais chez les plantes (Armbrust et al., 2004). L'étude du génome a également permis l'identification de gènes impliqués dans le transport et l'utilisation de l'acide silicique trouvé dans l'environnement pour la synthèse d'un squelette extérieur (exosquelette) composé de silice. Le génome de la diatomée *Phaeodactylum tricornutum* est d'une taille de 27,4 Mb (Bowler et al., 2008) pour 10402 gènes prédits. L'identification des gènes et l'analyse du génome ont été facilitées par la disponibilité de 130000 ESTs. Le génome de *P. tricornutum* présente une

structure relativement différente de celui de *T. pseudonana*. Aucune conservation majeure de la synténie n'a été observée entre les génomes des deux espèces. *P. tricornutum* partage 57% de ces gènes avec *T. pseudonana*. Parmi ceux-ci, 1328 gènes sont absents chez les autres eucaryotes séquencés et apparaissent donc spécifiques aux diatomées. Les auteurs ont également mis en évidence plusieurs centaines de gènes bactériens dont plus de 300 sont présents chez les deux diatomées. La plupart de ces gènes codent des composants susceptibles de fournir des capacités métaboliques originales, comme l'utilisation du carbone ou de l'azote organique. Ces gènes bactériens semblent dérivés d'origine multiple incluant les proteobactéries, les cyanobactéries et archae. Ce phénomène illustre un niveau sans précédent de transferts horizontaux de gènes entre les diatomées et les procaryotes, niveau largement supérieur à ce qui existe chez les autres eucaryotes séquencés (Bowler et al., 2008).

Disposer de génomes de différents straménopiles nous offrait donc une base de travail pouvant aider à l'annotation du génome de *Blastocystis* sp. De plus, ces organismes, pour lesquels nous disposons des génomes possèdent des modes de vie différents de celui de *Blastocystis* sp. puisqu'il s'agit de diatomées libres et de parasites de végétaux. Ceci apparaît d'autant plus intéressant que cela devrait permettre d'identifier des gènes spécifiques à ces modes de vies, notamment pour la mise en évidence de gènes spécifiques à *Blastocystis* sp. évoluant dans l'environnement intestinal des animaux. Dès lors, et afin d'identifier ces gènes, il était intéressant de pouvoir s'appuyer aussi sur des données génomiques connues chez d'autres protozoaires parasites vivant dans l'environnement intestinal et pour lesquels des gènes impliqués dans les phénomènes de pathogénie ont été identifiés. De manière plus généraliste, il était aussi intéressant de voir quel avait été l'apport de la génomique pour ces parasites.

3. Un exemple des intérêts de la génomique chez les parasites intestinaux, le cas des amibes du genre *Entamoeba*.

3.1 Caractéristiques et génomique des espèces du genre *Entamoeba*

Les amibes du genre *Entamoeba* sont des eucaryotes unicellulaires anaérobies, parasites obligatoires, vivant dans le tractus intestinal de l'homme et de divers animaux (Ximenez et al., 2009). Les espèces d'*Entamoeba* capables d'infecter l'homme sont l'espèce pathogène *E. histolytica*, et les espèces non pathogènes *E. dispar*, *E. moshkovskii*, *E. polecki*, *E. hartmanni* et *E. gingivalis*. Une autre espèce *E. invadens*, pathogène des reptiles, est utilisée comme modèle d'étude en raison de sa capacité à s'enkyster *in vitro* en culture axénique. 80 à 90% des individus infectés par *E. histolytica* sont asymptomatiques. Les premiers organes colonisés par le parasite sont la muqueuse intestinale du sigmoïde et le côlon, où seule l'espèce *E. histolytica*, responsable de la dysenterie amibienne et possédant un potentiel invasif, peut entraîner une dégradation des tissus. Cette espèce touche chaque année plus de 50 millions de personnes et serait responsable d'environ 100000 morts. La majorité des cas de mortalité est due à de sévères complications liées à l'envahissement intestinal ou extra-intestinal (Ximenez et al., 2009).

Trois génomes sont actuellement séquencés et annotés : *E. histolytica* souche HM-1 : IMMS, *E. dispar* souche SAW760 et *E. invadens* souche IP1 (Aurrecochea et al., 2011). Plusieurs autres souches sont en cours de séquençage. Le génome d'*E. histolytica* a une taille de 20 Mb pour 8201 gènes. Son génome a été annoté en 2005 (Loftus et al., 2005) avant d'être ré-assemblé et ré-annoté en 2010 (Lorenzi et al., 2010). *E. dispar*, morphologiquement identique à *E. histolytica* mais non pathogène, possède un génome de 30Mb pour environ 8700 gènes. *E. invadens* espèce pathogène des reptiles, provoquant notamment une mortalité significative chez les chelonians (tortues), possède un génome de 32Mb pour 11500 gènes (Aurrecochea et al., 2011).

3.2 L'analyse du génome d'*E. histolytica*, un apport de nouvelles données

L'analyse du génome d'*E. histolytica* (Loftus et al., 2005; Lorenzi et al., 2010) a mis en évidence une grande variété de métabolismes partagés avec d'autres protistes pathogènes : *Giardia lamblia* et *Trichomonas vaginalis*. Ces adaptations incluent notamment la perte ou la réduction de voies métaboliques mitochondriales et la présence d'enzymes liées au stress oxydatif. Les différentes études phylogénétiques menées ont également mis en évidence des transferts latéraux, ce qui a eu pour effet d'augmenter le répertoire métabolique. En effet, le métabolisme d'*E. histolytica* semble avoir été formé par des pertes secondaires de gènes et des transferts horizontaux de gènes principalement bactériens. En l'absence de certaines protéines du cycle de Krebs et de la chaîne de phosphorylation oxydative, l'amibe fait de la fermentation. Ainsi chez *E. histolytica*, la glycolyse fournit le pyruvate qui est converti en acétyl-coA sous l'action d'une Pyruvate : ferrédoxine oxydoréductase; l'acétyl-coA serait lui transformé en acétate avec la génération d'ATP ou réduit en éthanol avec régénération de NAD. L'analyse du génome a permis de mettre en évidence une glycolyse incomplète avec quelques enzymes impliquées, notamment une phosphoglucose isomérase, une phosphoglycérate kinase, et une phosphoglycérate mutase (Anderson et Loftus, 2005). *E. histolytica* a également perdu les voies de synthèse des acides aminés à l'exception de la voie de synthèse des sérines et cystéines, probablement dû à son statut de résident du côlon humain et donc à un accès facile à des bactéries et composés organiques pouvant être phagocytés, lui fournissant ainsi les apports en acides aminés nécessaires à son développement. Le maintien de ces voies métaboliques et d'un niveau élevé de cystéines a probablement pour objectif de compenser l'absence de glutathion et des enzymes associées, qui sont des composants majeurs dans la résistance au stress oxydatif. De même *E. histolytica* est incapable de synthétiser des acides gras mais a conservé la capacité à synthétiser des phospholipides. Tous ces gènes et l'existence de voies métaboliques originales, sont des cibles à étudier pour le développement de nouvelles drogues.

L'étude de son génome a également permis de mettre en évidence plusieurs facteurs de pathogénie parmi lesquels (1) des lectines, impliquées dans l'adhésion aux cellules de l'hôte, (2) des protéases à cystéine impliquées dans la dégradation de la matrice extracellulaire de l'hôte et du mucus, et connues pour être impliquées dans la virulence, et (3) des peptides formant des pores (amoebapores) capables de lyser des cellules (Loftus et al., 2005; Clark et al., 2007). L'analyse du génome a montré une redondance dans les gènes codant ces facteurs de pathogénie. Dix nouvelles protéases à cystéine, avec un signal N-terminal d'ancrage transmembranaire prédit, laissant supposer une localisation à la surface de l'amibe, ont été identifiées. En plus de 3 nouvelles amoebapores, une protéine homologue de l'hémolysine a été identifiée, laissant supposer que cette dernière pouvait être impliquée dans la lyse. Tous ces gènes identifiés chez *E. histolytica* sont d'intérêt pour comprendre l'adaptation à l'environnement intestinal et sont autant de candidats à rechercher chez *Blastocystis* spp. dans le but de révéler son pouvoir pathogène.

3.3 *E. invadens* et *E. histolytica*, deux amibes pathogènes relativement proches

E. invadens, parasite de reptiles, est susceptible de devenir pathogène et d'entraîner des pathologies au niveau du côlon similaires à celles provoquées par *E. histolytica* chez l'homme. Elle est utilisée comme amibe modèle pour l'étude de l'enkystement. Globalement, le génome d'*E. invadens* possède plusieurs caractéristiques identiques au génome d'*E. histolytica* (Wang et al., 2003). Ainsi les séquences hautement répétées d'*E. histolytica* ont également été trouvées chez *E. invadens*. De plus, *E. invadens* possède plusieurs protéines homologues à des protéines connues pour être impliquées dans la virulence (lectines, amoebapores, protéases à cystéine) chez *E. histolytica*. Le séquençage a ainsi permis de mettre en évidence 12 nouvelles protéases à cystéine, portant leur nombre connu à 20 chez *E. invadens*, et dont 14 sont homologues à celles identifiées chez *E. histolytica*. De même *E. invadens* possède cinq séquences similaires aux amoebapores connues chez *E. histolytica* et *E. dispar*. Les alignements multiples de

ces séquences montrent 2 séquences homologues aux amoebapores A et 2 aux amoebapores B. Les 2 peptides homologues aux amoebapores B possèdent une forte similarité entre eux, laissant suggérer une récente duplication. Au niveau du métabolisme, *E. invadens* possède des enzymes impliquées dans la fermentation, présentant de fortes identités avec leurs homologues chez *E. histolytica* qui pour la plupart de ces enzymes semblent issues de transferts latéraux de procaryotes anaérobies. La forte similarité de ces enzymes entre *E. invadens* et *E. histolytica* suggère que les transferts latéraux entre des procaryotes anaérobies et le genre *Entamoeba* se sont déroulés bien avant la séparation entre *E. invadens* et *E. histolytica*. Une différence demeure toutefois entre les 2 espèces, les enzymes fermentatives sont présentes en plusieurs copies chez *E. invadens*, alors qu'elles sont en copie unique chez *E. histolytica*. Cette duplication spécifique à l'espèce *E. invadens* reste peu claire. Toutes ces données ont permis de mettre en évidence un bon nombre d'éléments reflétant le mode de vie parasitaire d'*E. invadens*, et la présence de gènes homologues avec l'espèce pathogène *E. histolytica*.

3.4 *Entamoeba dispar*, une amibe non pathogène utile pour la génomique comparative

E. dispar est morphologiquement identique à *E. histolytica*, mais a un potentiel virulent très différent. *E. histolytica*, est capable d'être invasive alors qu'*E. dispar*, si elle est capable de coloniser l'intestin de l'homme ne possède aucun potentiel invasif. Quelques différences clés concernant la virulence entre les 2 espèces ont été identifiées, notamment au niveau des protéases à cystéine. Ainsi, certaines protéases à cystéine sont absentes ou dégénérées chez *E. dispar*. D'autres, conservées chez les deux espèces, présentent des différences d'expression avec un faible niveau d'expression chez *E. dispar* (Bruchhaus et al., 1996; Bruchhaus et al., 2003). Cependant les différences essentielles entre les souches virulentes et les non virulentes, déterminant le pouvoir invasif, restent méconnues. Une étude sur la différence génétique entre les deux espèces par puces à

ADN a mis en évidence des régions génétiques divergentes entre *E. histolytica* et *E. dispar* (Shah et al., 2005). La comparaison de ces régions génétiques a permis d'identifier des régions utilisables pour typer les souches ou les espèces, ainsi que 67 gènes fortement ou significativement différents entre les deux (Shah et al., 2005). Pourtant si ces différences contribuent à la virulence, les phénotypes restent à déterminer. Une autre approche par puces à ADN, consistant à étudier les profils d'expression de 2110 gènes uniques, entre la souche virulente *E. histolytica* HM-1 : ISS et les souches non virulentes *E. histolytica* Rahman et *E. dispar* SAW760 a été réalisée en 2006 (MacFarlane et Singh, 2006). Par cette technique, 415 gènes apparaissent plus faiblement exprimés chez *E. dispar*, et 32 gènes chez *E. histolytica* Rahman, par rapport à *E. histolytica* HM-1 : ISS. Plus intéressant, 29 gènes ont une expression diminuée à la fois chez *E. dispar* et *E. histolytica* Rahman. La plupart des gènes de cette catégorie sont des gènes codant des protéines hypothétiques, mais certains codent des protéines de stress HSP associées au réticulum endoplasmique, des protéines de surface ou encore une sérine palmitoyltransférase (synthèse de sphingolipides). L'expression de ces gènes semble être corrélée avec la virulence. Ils pourraient donc jouer un rôle dans la pathogénie des amibes.

Ces approches ayant permis de mettre en évidence des facteurs de pathogénie chez les différentes espèces d'*Entamoeba* sont envisageables chez *Blastocystis* spp. Celles-ci sont d'autant plus intéressantes puisque la pathogénie de *Blastocystis* spp. est supposée être dépendante des sous-types.

L'analyse des génomes des différentes espèces d'*Entamoeba* a apporté un nombre important de données sur la biologie de ces organismes, leurs métabolismes mais également leur pathogénie. Ces analyses ont notamment permis de mettre en évidence de nombreux facteurs de virulences chez *E. histolytica* : des lectines, des amoebapores et les protéines associées, mais surtout un nombre considérable d'endopeptidases et notamment de protéases à cystéine (Clark et al., 2007). L'élucidation du rôle précis de

ces enzymes est un challenge pouvant largement aidé à la compréhension des mécanismes de virulence. L'analyse du génome d'*E. histolytica* a aussi mis en évidence un tiers de gènes sans homologue chez les eucaryotes supérieurs, qui sont de fait spécifiques à *Entamoeba* spp. Ces gènes ainsi que ceux codant les protéases à cystéines constituent des cibles de choix pour le développement de nouvelles drogues antiparasitaires.

III. Les protéases, des enzymes essentielles dans la biologie des parasites

Tous les parasites doivent infecter leur hôte pour pouvoir survivre, se développer et se multiplier. Ces derniers sont susceptibles d'entraîner des modifications structurelles et fonctionnelles au niveau des tissus ou de la physiologie de l'hôte. En réponse à ces agressions, l'hôte a développé une série de mécanismes de défense pour limiter la population parasitaire et contrebalancer les effets physiopathologiques. Ces mécanismes peuvent être regardés comme un processus d'adaptation au parasitisme. A l'instar de l'hôte, les parasites ont développé diverses stratégies complexes d'adaptation ou d'échappement à la réponse immunitaire afin de persister dans un environnement qui leur est souvent hostile. Les protéases jouent des rôles-clés au sein de ces processus en intervenant à différents niveaux de l'interaction hôte-parasite. Elles facilitent ainsi la pénétration du parasite dans les tissus de l'hôte, jouent un rôle dans la nutrition à travers la dégradation des protéines de l'hôte entraînant des mécanismes pathologiques pour l'hôte, et participent à l'échappement vis-à-vis du système immunitaire de l'hôte. Ces différents rôles remplis par les protéases en font des cibles potentielles en thérapie et vaccination antiparasitaire.

1. Les protéases, un vaste ensemble

Les protéases catalysent le clivage de liaisons peptidiques au sein de peptides et de protéines. Ce clivage peut se faire au niveau d'une liaison peptidique interne, on parle d'endopeptidases ou au niveau d'acides aminés situés aux extrémités N-terminale ou C-terminale, on parle alors d'exopeptidases ou plus spécifiquement d'aminopeptidases et de carboxypeptidases.

Historiquement, les protéases ont été subdivisées en quatre grandes classes : les protéases à serine, les protéases à cystéine, les protéases à aspartate, et les métalloprotéases. Cette classification est fonction des acides aminés du site actif

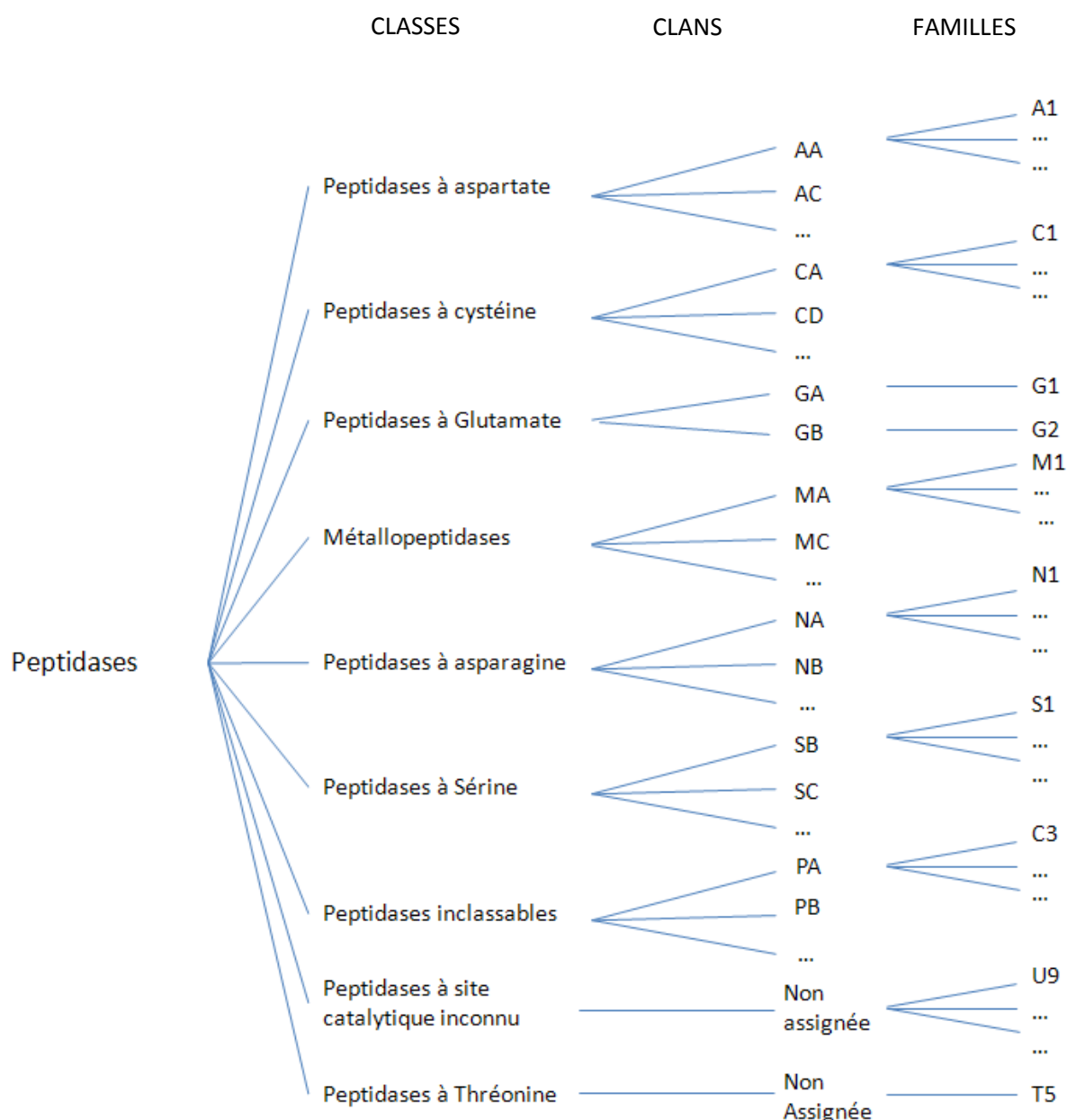


Figure 13 : Classification hiérarchique des peptidases définie par la base de données MEROPS. Les peptidases sont assignées à une famille constituant un groupe présentant des homologies de séquences significatives. Les familles sont groupées dans des clans. Les membres des différents clans ne présentent pas d'homologies significatives entre eux (d'après Atkinson et al., 2009).

impliqués dans la liaison du substrat, et indispensable à la catalyse de la réaction, ou dans le cas des métalloprotéases en raison de l'ion situé au niveau du site actif (McDonald, 1985).

Cette classification fournit des indices sur la fonction *in vivo* de ces enzymes (McKerrow et al., 1993). En 1996, la base de données MEROPS a été créée pour répertorier et classer toutes les protéases. Celle-ci a une classification hiérarchique dans laquelle chaque protéase appartient à une famille et à un clan regroupant les familles homologues sur la base d'identités de séquence et d'un ancêtre commun (Figure 13). Chaque peptidase, famille et clan ont un identifiant unique (Rawlings et al., 2006; Rawlings et al., 2008; Rawlings et al., 2010). Pour chaque peptidase, plusieurs informations vont être disponibles, l'architecture de son domaine, sa classification MEROPS, une description de son activité, de ses fonctions biologiques, et une spécification des sites de clivage. La dernière mise à jour de MEROPS (release 9.05) fait mention de 7 classes de peptidases : 1. Peptidase à aspartate, 2. Peptidase à cystéine, 3. Peptidase à glutamate, 4. Métallopeptidase, 5. Peptidase à asparagine, 6. Peptidase à sérine, 7. Peptidase à thréonine. A ces classes viennent s'ajouter, une classe regroupant les peptidases à site catalytique inconnu, et une classe regroupant les peptidases inclassables. Elles sont classées en 226 familles pour 57 clans.

1.1 Les quatre grandes classes de peptidases

1.1.1 Les protéases à sérine

Le terme protéase à sérine indique une classe de protéases qui utilisent un résidu sérine pour leur réaction catalytique. Elles sont impliquées dans une large variété de processus physiologiques, et sont largement répandues dans la nature, dans tous les règnes du vivant y compris chez les virus. Chez l'homme par exemple, ces enzymes interviennent au niveau de la digestion, mais également dans la coagulation sanguine, et la lutte contre les infections. Chez les parasites elles interviennent dans la régulation de

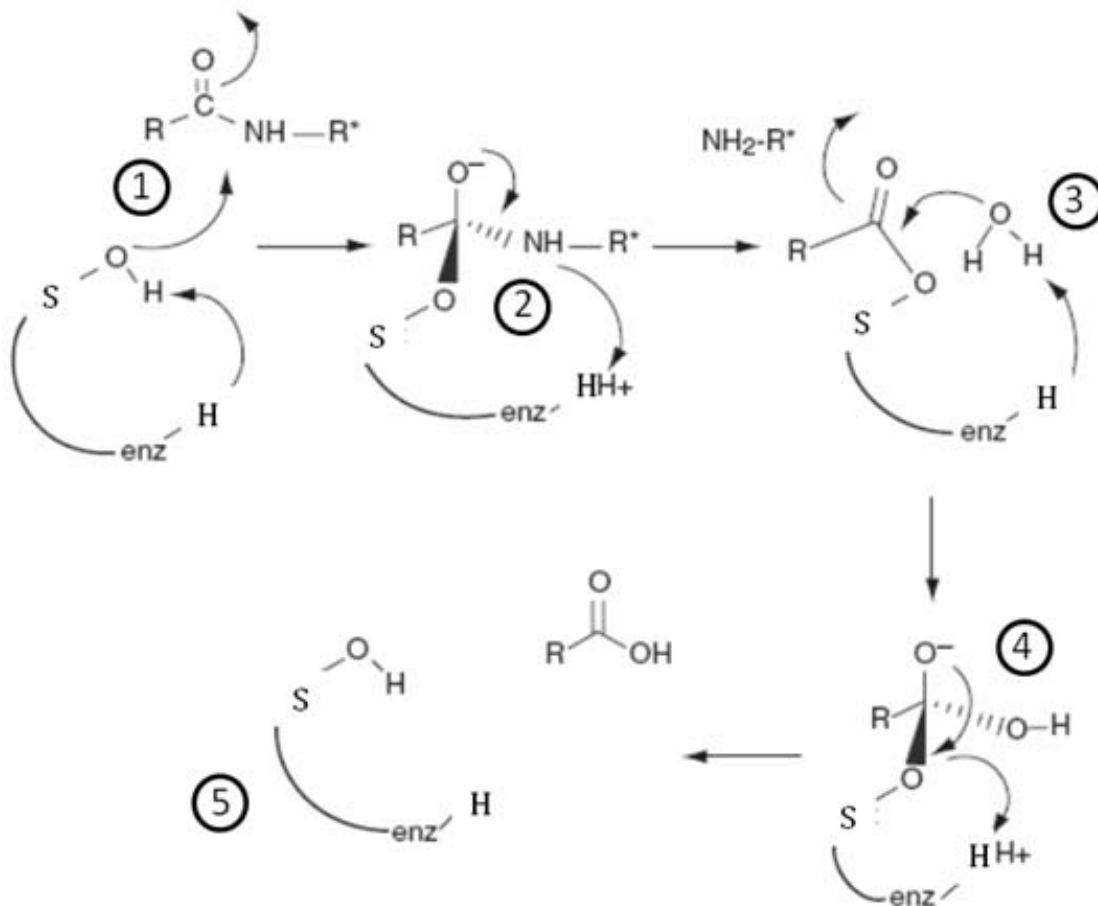


Figure 14: Schéma du mécanisme d'action des protéases à sérine (d'après Hedstrom 2002). 1- Attaque du groupement carbonyle de la liaison peptidique par l'hydroxyle de la sérine entrainant la formation d'un intermédiaire tétraédrique, l'histidine jouant le rôle d'accepteur de protons. 2- Transfert d'un proton de l'histidine au groupement amine de la liaison peptidique, entrainant la libération du résidu peptidique et la formation d'un acylenzyme intermédiaire. 3- Attaque de l'acylenzyme par une molécule d'eau qui va être déprotonée entrainant la formation d'un deuxième intermédiaire tétraédrique, l'histidine jouant le rôle d'accepteur de protons. 4- Protonation de la sérine. 5- Retour à l'état initial de l'enzyme et libération du deuxième résidu peptidique.

leur cycle de développement mais également dans l'infection des cellules et des tissus de l'hôte.

Actuellement plus de 74000 séquences de protéases à sérine sont classées dans la base de données MEROPS et regroupées dans 15 clans et 49 familles. Il existe des différences significatives dans la distribution des protéases à sérine dans les clans et entre les espèces. La plupart des protéases à sérine sont des endopeptidases, même si quelques familles d'exopeptidases ont été décrites (Page et Di Cera, 2008). Une des caractéristiques des protéases à sérine est de posséder une triade d'acides aminés Ser-His-Asp au niveau de son site actif. Le clivage des liaisons peptidiques par les protéases à sérine va impliquer le groupement hydroxyle de la sérine. Celui-ci va servir de catalyseur, en jouant le rôle de nucléophile, et lier le carbonyle de la liaison peptidique (Figure 14). L'histidine agit alors comme accepteur de protons. L'aspartate joue elle un rôle de maintien du réseau hydrogène établi entre l'histidine et la sérine. Le résultat de cette étape dite d'acylation est la formation d'un intermédiaire tétraédrique. Il y a alors transfert du proton de l'histidine au groupe amine de la liaison peptidique entraînant la libération du résidu peptidique, et aboutissant à un acylenzyme intermédiaire. Cet acylenzyme va être attaqué par une molécule qui va être déprotonée formant ainsi un deuxième intermédiaire tétraédrique. La protonation de la sérine va entraîner un retour à l'état initial de l'enzyme et la libération de l'autre résidu peptidique (Hedstrom, 2002). Si la triade d'acides aminés est retrouvée dans beaucoup de protéases à sérine, quelques peptidases utilisent un mécanisme composé uniquement de deux acides aminés avec une lysine ou une histidine en combinaison avec la sérine. D'autres utilisent une triade composée de résidus différents, comme une paire d'histidine combinée à la sérine (Page et Di Cera, 2008).

1.1.2 Les protéases à cystéine

Les protéases à cystéine sont définies ainsi en raison de l'utilisation du groupement thiol d'un résidu cystéine comme nucléophile lors de la catalyse. La toute

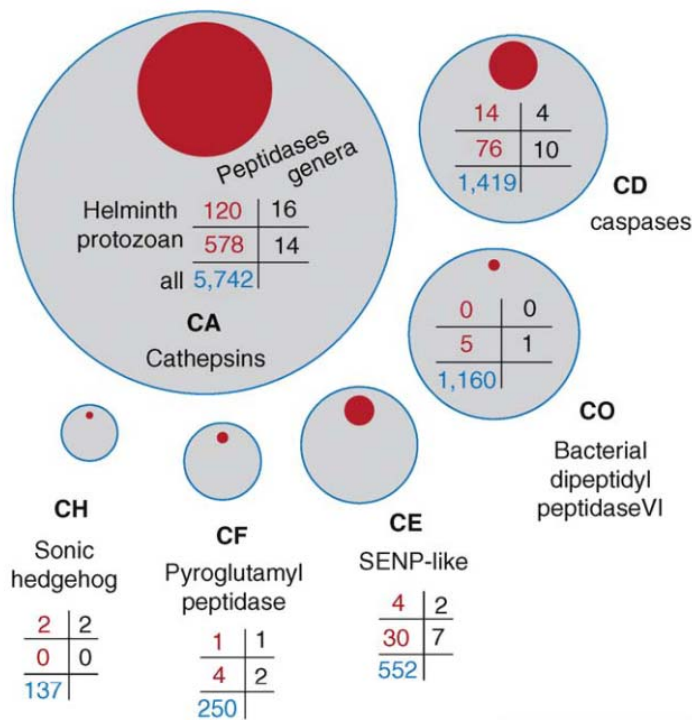


Figure 15: Vue d'ensemble des 6 clans contenant des protéases à cystéine parasitaires. Les protéases à cystéine parasitaires sont réparties au sein des clans CA, CD, CE, CF, CH et CO. Les ronds rouges représentent les protéases à cystéine des protozoaires et des helminthes parasites. Leur taille est proportionnelle à la population totale de chaque clan. Les ronds gris représentent la population totale de chaque clan, leur taille est proportionnelle à la population totale de toutes les peptidases à cystéine des différents clans (Atkinson et al 2009).

première protéase à cystéine, la papaïne, fut caractérisée et purifiée en 1879, à partir de la papaye. On retrouve des protéases à cystéine partout dans le règne vivant (virus, bactéries, champignons, protozoaires, mammifères, plantes) où elles jouent des rôles importants dans de multiples processus biologiques. A titre d'exemple, les protéases à cystéine des mammifères sont impliquées dans le remodelage des tissus, le renouvellement des matrices extracellulaires, les fonctions du système immunitaire, mais également dans les processus de l'inflammation, de la présentation de l'antigène, dans l'apoptose ou encore dans les cancers (Dickinson, 2002). Elles sont capables d'influencer les processus cellulaires comme la différenciation, sont impliquées dans la signalisation cellulaire et l'activation de récepteurs spécifiques ou de facteurs de croissance (Obermajer et al., 2008). Chez les plantes, leurs fonctions sont tout aussi vastes (Grudkowska et Zagdanska, 2004) puisqu'elles sont impliquées dans la croissance et le développement. Elles interviennent dans la dégradation des protéines de stockage, dans les réponses aux stress biotiques et abiotiques, dans la sénescence des organes et les processus de mort cellulaire programmée. Chez les parasites, elles jouent des rôles-clés dans la virulence, l'invasion cellulaire et tissulaire, l'évasion du système immunitaire, l'enkystement et le dékystement mais également dans la régulation du cycle de développement des parasites. Leurs implications dans ces différents processus en font des cibles intéressantes pour l'établissement de molécules ou de vaccins antiparasitaires (Sajid et McKerrow, 2002).

Les fonctions-clés remplies par cette classe de protéases dans la biologie des organismes font que la régulation de leurs activités est essentielle pour les cellules et les tissus. En effet, une activité protéolytique au mauvais endroit et/ou au mauvais moment pourrait entraîner des dommages sur les tissus ou l'organisme, dommage pouvant entraîner la létalité. (Wiederanders et al., 2003). Beaucoup de protéases à cystéine sont ainsi synthétisées sous une forme immature avec une pro-région clivable et un domaine mature possédant l'activité catalytique. Cette pro-région possède plusieurs fonctions indépendantes. Elle est impliquée dans la régulation de l'activité protéasique. Elle sert

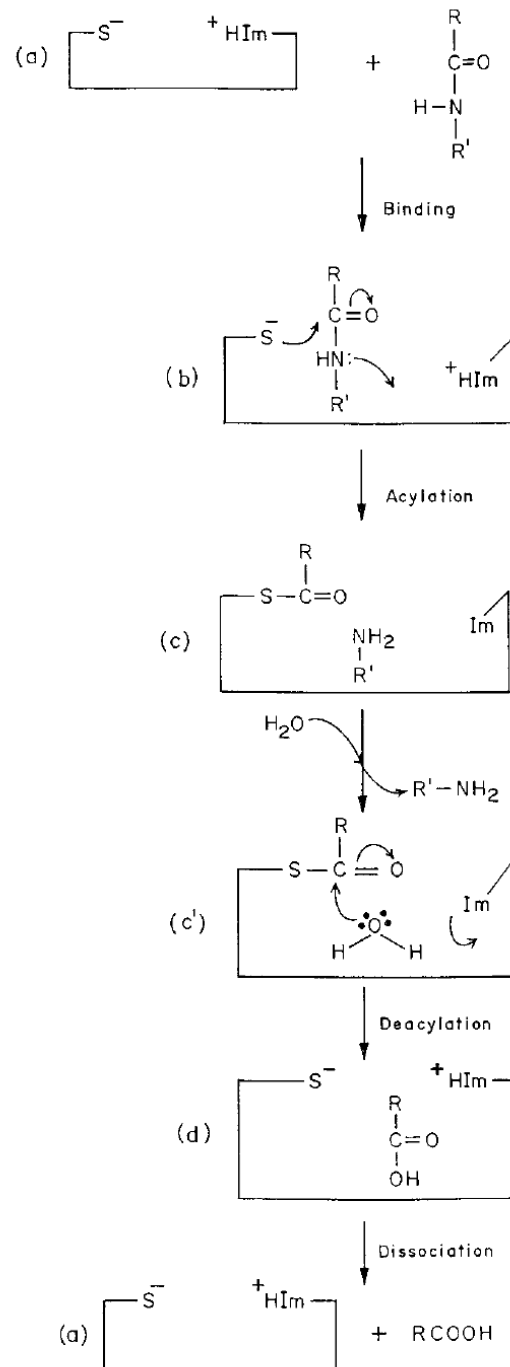


Figure 16 : Schéma du mécanisme d'action des cystéines protéases. L'étape initiale (a) implique la liaison non covalente de l'enzyme libre et du substrat pour former un complexe (b). Il s'ensuit l'acylation de l'enzyme (c) avec la formation et la libération d'un produit amine $\text{R}'\text{-NH}_2$. L'étape suivante est la déacylation de l'enzyme (c') qui va réagir avec une molécule d'eau et libérer un second produit (d) avec la régénération de l'enzyme libre (Rao et al., 1998).

aussi de chaperonne, participe au repliement des protéines, et dans quelques cas, cette région sert de signal pour la localisation intracellulaire des protéases (Sajid et McKerrow, 2002).

Actuellement, 32794 séquences de protéases à cystéine sont classées dans la base de données MEROPS. Elles sont regroupées en 12 clans et 70 familles. Certaines familles de protéases à cystéine n'ont pas de clans assignés. Les protéases à cystéine des parasites ne sont retrouvées que dans 6 clans : CA, CD, CE, CF, CO et PD (Figure 15). La majorité de ces peptidases appartiennent au clan A (plus de 84% des protéases à cystéine connues), parmi lesquelles on retrouve les cathepsines, et au clan D qui regroupe notamment les caspases et les légumaines (Atkinson et al., 2009). Le mécanisme catalytique des protéases à cystéine est similaire aux protéases à sérine (Figure 16). La différence principale concerne le nucléophile qui va être joué par le groupement thiol de la cystéine. Il est fréquent qu'un résidu asparagine soit impliqué pour stabiliser les liaisons. A l'image des protéases à sérine, les acides aminés du site catalytique varient. Dans la majorité des protéases du clan A, on retrouve les acides aminés Cys, His et Asn impliqués dans la catalyse. On ne retrouve que la cystéine et l'histidine pour celles du clan B, et celles du clan CM possèdent un site actif constitué d'une triade His-Glu-Cys. Ceci traduit la variabilité des sites actifs au sein d'une même classe d'enzyme (Rudenskaya et Pupov, 2008).

Actuellement, si les protéases à cystéine les plus étudiées et les plus répandues sont des protéases de plantes, de plus en plus sont identifiées chez les microorganismes pathogènes où ces enzymes apparaissent comme des facteurs de virulence.

1.1.3 Les métalloprotéases

Les métalloprotéases sont un groupe très hétérologue d'enzymes protéolytiques dont la caractéristique principale est de contenir un ou deux ions métalliques dans leur site actif. La majorité des métalloprotéases ont un ion Zn^{2+} , quelques unes ont un ion Co^{2+} , Mg^{2+} , Ni^{2+} , ou Cu^{2+} . Cet ion métallique intervient pour activer une molécule d'eau

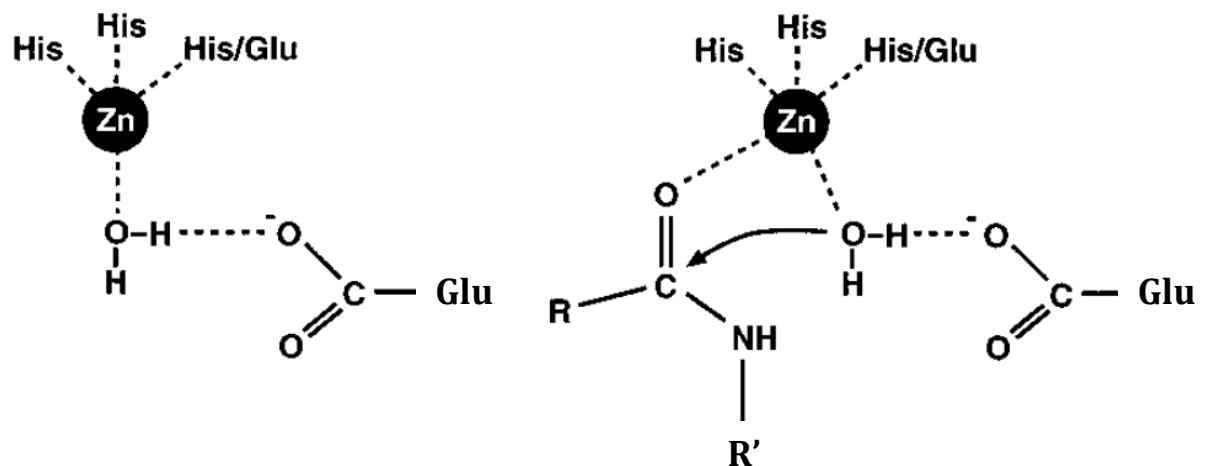


Figure 17 : Schéma simplifié du mécanisme d'action des métalloprotéases (d'après Lipscomb and Strater, 1996). Attaque nucléophile par une molécule d'eau du carbone du groupement carbonyle de la liaison peptidique. Cette attaque est favorisée par l'ion zinc et un acide aminé glutamate. Elle va aboutir à la libération d'un produit amine R'-NH₂ et d'un produit carbonyle R-COOH.

qui servira de nucléophile dans la catalyse de la réaction (Rawlings et al., 2004). Ces enzymes sont représentées chez les eucaryotes, aussi bien chez les espèces plantes, animales que les microorganismes et les procaryotes. De plus, elles présentent de fortes similarités en termes de structures et de propriétés (Rawlings et Barrett, 1995).

Certaines métalloprotéases fonctionnent dans la cellule ou sur la membrane cellulaire. Les autres sont sécrétées à l'extérieur de la cellule. Les métalloprotéases extracellulaires de mammifères sont très étudiées car impliquées dans de nombreuses maladies comme l'arthrite, les maladies du système nerveux central, ou encore dans certaines tumeurs malignes (Malemud, 2006). Les bactéries sécrètent des métalloprotéases pour dégrader les protéines de leur environnement à des fins nutritives. Ces métalloprotéases sont connues pour être des facteurs de virulence importants des bactéries, comme la thermolysine chez *Bacillus thermoproteolyticus*, l'aureolysine chez *Staphylococcus aureus* ou encore une des trois protéines de la toxine de l'anthrax chez *Bacillus anthracis* (Wu et Chen, 2011). Chez les protozoaires parasites (*Leishmania*, *trypanosoma*), les métalloprotéases ont été rapportées comme facteurs corrélés à la virulence. Chez les apicomplexes (*Plasmodium*, *Eimeria*, *Cryptosporidium*), leur implication dans l'invasion et la dissémination parasitaire a été explorée (Villena et al., 2006). Chez les leishmanies, McGwire a montré *in vitro* que la migration du parasite à travers la matrice extracellulaire était augmentée par la glycoprotéine gp63, qui agit en dégradant les composants de cette matrice, suggérant un rôle important de cette métalloprotéase dans la pathogénie (McGwire et al., 2003).

Le mécanisme d'action des métalloprotéases est légèrement différent des autres protéases. Il dépend de la présence d'un ou deux cations bivalents, d'une molécule d'eau et d'un motif de liaison à l'ion métallique (Lipscomb et Strater, 1996)(Figure 17). La majorité des familles des métalloprotéases possède un motif pentapeptide His-Glu-Xaa-Xaa-His (HEXXH) ou les histidines servent de ligand, avec la présence d'un autre acide aminé qui vient compléter le site de liaison. Dans le cas de la thermolysine, il a été suggéré qu'un acide aminé Glu assiste l'attaque nucléophile d'une molécule d'eau sur le

carbone de la fonction carbonyle de la liaison peptidique, qui est polarisé par l'ion Zn^{2+} . Les familles sont regroupées dans les différents clans selon la nature des acides aminés qui vont compléter le site de liaison à l'ion métallique. Par exemple, le clan MA a la séquence HEXXH-E, le clan MB possède lui la séquence HEXXH-H (Rao et al., 1998). Il y a à l'heure actuelle plus de 63900 séquences de métalloprotéases dans la base de données MEROPS, regroupées en 14 clans et 62 familles. 8 familles ne sont pas assignées à un clan.

1.1.4 Les protéases à aspartate ou protéases acides

Les protéases à aspartate sont une classe d'enzymes très largement répandue dans le monde vivant. On les retrouve chez les vertébrés, les plantes, les levures, les champignons, les parasites ou les virus. Dans la base de données MEROPS, elles sont classées en 16 familles, réparties dans 5 clans, avec 2 familles n'ayant pas de clan assigné. Elles sont caractérisées par la présence de deux résidus Asp responsables de l'activité catalytique, une structure tridimensionnelle très similaire entre elles, et une activité à pH acide (Davies, 1990). Cependant si la majorité des protéases à aspartate sont conformes à ces caractéristiques, elles diffèrent substantiellement en termes de propriétés catalytiques, de localisation cellulaire et de fonctions biologiques.

La majorité des protéases à aspartate est synthétisée sous forme de zymogènes, précurseurs inactifs des enzymes, ce qui leur confère une protection contre la lyse. Le zymogène est converti en enzyme active par le clivage d'une région en N-terminal appelée pro-peptide, région d'environ 50 acides aminés et riche en résidus basiques. Ce pro-peptide sert à stabiliser la forme inactive et prévient l'entrée du substrat dans le site actif. La conversion de la forme inactive à la forme active reste un processus complexe. En raison des différences dans les mécanismes d'activation des protéases à aspartate, trois mécanismes d'activation ont été proposés. Un mécanisme d'auto-maturation complet, un mécanisme de maturation impliquant d'autres enzymes, et un mécanisme totalement réalisé par d'autres enzymes (Horimoto et al., 2009). Concernant leurs

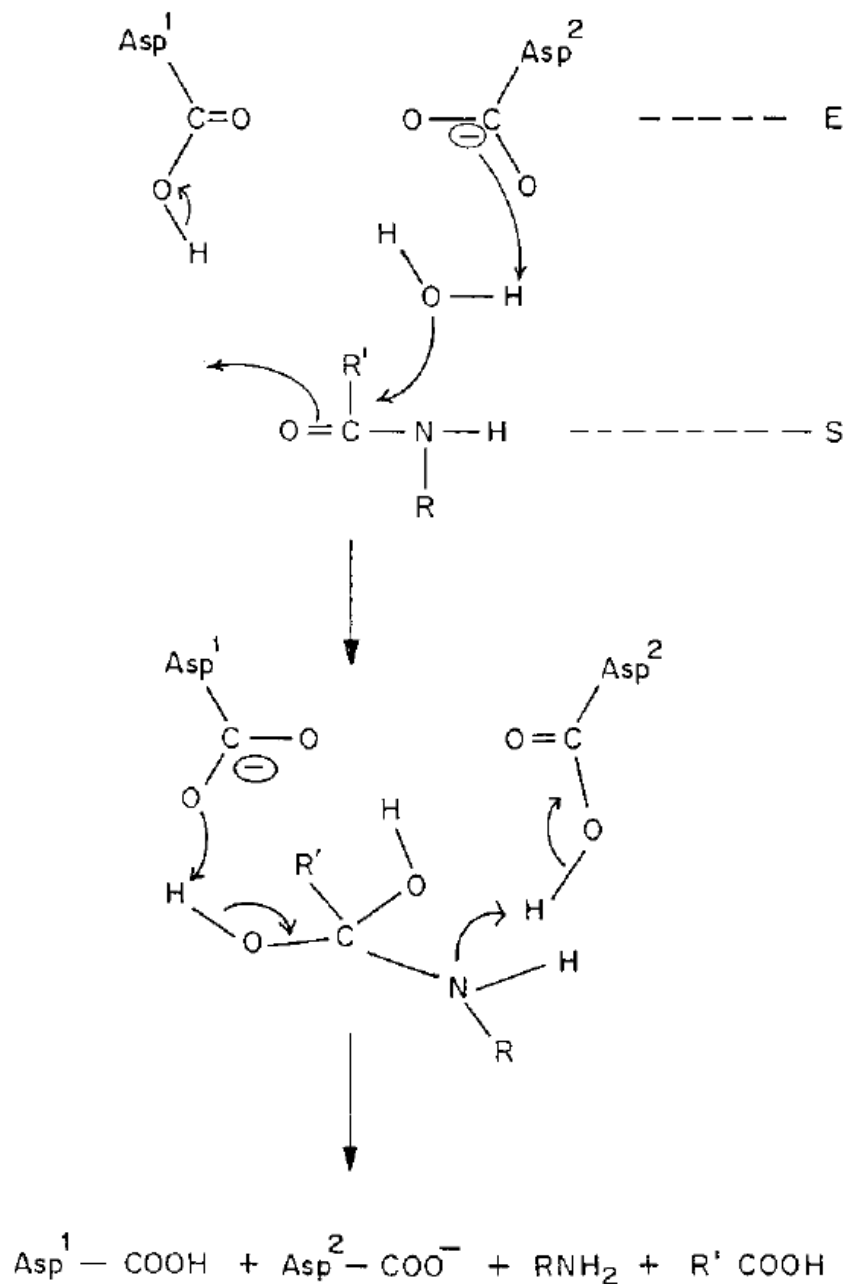


Figure 18 : schéma simplifié du mécanisme d'action des protéases à aspartate (Rao et al., 1998). Attaque nucléophile de la liaison peptidique par une molécule d'eau suite à son activation par un résidu aspartate entraînant la formation d'un carbone tétraédrique. L'autre résidu aspartate se déprotonne pour stabiliser cet état dit intermédiaire. Un proton est alors transféré sur l'azote de la liaison peptidique, entraînant la décomposition du carbone tétraédrique et la rupture de la liaison peptidique donnant ainsi les différents produits de la réaction.

fonctions biologiques, elles varient selon les espèces. Chez les plantes elles sont impliquées (1) dans la dégradation des protéines à des fins de source d'azote, (2) dans la réponse au stress à travers la dégradation de protéines impliquées dans la pathogénie, ou encore (3) dans la sénescence et la mort cellulaire programmée (Cooper, 2002). Chez les animaux, on retrouve la rénine qui a une action hypertensive à travers son rôle dans le système rénine-angiotensine. Chez l'homme, des protéases à aspartate ont été montrées impliquées dans la maladie d'Alzheimer, les ulcères gastriques ou encore le cancer du sein (Coombs et al., 2001). Chez les parasites, les protéases à aspartate sont impliquées dans les phénomènes de pathogénie et sont largement ciblées par les molécules antiparasitaires. *Plasmodium falciparum*, par exemple, dégrade l'hémoglobine de son hôte pour fournir les nutriments nécessaires à sa croissance et sa maturation. Ce processus se déroule dans la vacuole alimentaire qui contient notamment des protéases à aspartate (plasmepsines), des protéases à cystéine et des métalloprotéases, toutes impliquées dans la voie de dégradation de l'hémoglobine. La dégradation de l'hémoglobine étant vitale pour le parasite, le blocage de cette voie offre une approche valide pour le développement d'agents chimiothérapeutiques. En culture, le blocage des protéases à aspartate par des inhibiteurs sélectifs prévient la dégradation de l'hémoglobine entraînant ainsi une mort du parasite (Sajid et McKerrow, 2002).

Le mécanisme d'action de ces enzymes implique la coordination d'une molécule d'eau et des deux résidus Asp. Un Asp va activer la molécule d'eau en soustrayant un proton, permettant ainsi à la molécule d'eau d'attaquer le carbone carbonyle de la liaison peptidique du substrat (Figure 18).

Les protéases, autrement dit les enzymes assurant l'hydrolyse d'autres protéines, jouent donc des rôles importants dans de nombreux processus physiologiques et sont des composants critiques pour toutes les cellules vivantes. Leurs différents rôles relatifs notamment à leurs implications dans des pathologies en ont fait des sujets d'études intéressants. Leurs multiples propriétés font qu'elles sont maintenant utilisées en médecine pour leurs effets anti-inflammatoires et leurs rôles en tant que facteurs de

virulence chez certaines espèces pathogènes en ont fait des cibles de choix à des fins thérapeutiques. A titre d'exemple des inhibiteurs de protéases sont actuellement utilisés pour bloquer la réplication du VIH et du virus de l'hépatite C. A l'heure actuelle de nombreux travaux portent sur l'utilisation d'inhibiteurs de protéases pour essayer de traiter de nombreuses maladies virales, cancéreuses ou provoquées par des parasites.

2. Rôle des protéases chez les protozoaires parasites

2.1 Quelques exemples de l'implication des protéases dans les phénomènes de pathogénie

2.1.1 Implication dans l'adhésion des parasites aux épithéliums

L'adhésion à l'épithélium de l'hôte est une étape critique dans le développement des parasites extracellulaires. En effet, pour pouvoir envahir un tissu ou simplement se maintenir dans un environnement il leur est nécessaire de se lier aux cellules. Pour les parasites intestinaux, par exemple, cette adhésion aux cellules de l'épithélium intestinal va leur permettre d'obtenir les nutriments nécessaires à leur survie, par phagocytose de bactéries ou de débris intestinaux. Cette étape est également essentielle pour la colonisation de l'intestin et pour permettre la pénétration des tissus de l'hôte. Cette étape d'adhésion est d'autant plus difficile que le tractus gastro-intestinal des mammifères est revêtu d'une couche protectrice de mucus qui est la première couche de défense contre l'invasion. Il constitue une barrière physique entre le tissu sous-jacent et les micro-organismes. Il prévient leur adhésion en les piégeant et en facilitant leur expulsion. Pour persister les parasites ont donc développé des mécanismes d'adhésion. Chez l'amibe *Entamoeba histolytica*, le complexe EhCPADH, impliqué dans l'adhésion du parasite, est un hétérodimère formé par une adhésine EhaDH112 et une protéase à cystéine EhCP112. L'utilisation d'anticorps contre ce complexe entraîne une diminution de plus de 50% de l'adhésion aux cellules-hôtes, suivie par une diminution de la phagocytose et de la destruction des cellules de l'hôte par *E. histolytica* (Sateriale et

Huston, 2011). La protéase à cystéine EhCP112 est retrouvée faiblement exprimée chez des mutants avirulents (Ocadiz et al., 2005). En revanche, chez le type sauvage, cette protéine est localisée dans des vésicules de sécrétion. La protéine recombinante est capable de perturber une monocouche cellulaire de cellules de rein de chien (MDCK) et de digérer les protéines de la matrice extracellulaire. Cette protéine est sécrétée par les stades trophozoïtes du parasite, stades susceptibles d'envahir la muqueuse digestive.

L'implication des protéases dans l'adhésion aux cellules d'un épithélium peut également être illustrée par le parasite *Trichomonas vaginalis*. *T. vaginalis* est un protozoaire flagellé responsable de la trichomonase, une des maladies sexuellement transmissibles les plus répandues dans le monde. Il infecte l'épithélium du tractus génital et peut causer des vaginites, des cervicites, ou des urétrites, mais peut aussi, en cas d'infection chronique, favoriser le développement de cancer ou provoquer des avortements spontanés chez les femmes enceintes. L'adhésion du parasite aux cellules épithéliales est hautement spécifique. C'est une étape critique dans l'initiation de l'infection. *T. vaginalis* possède un niveau élevé d'activité protéolytique, principalement de type protéase à cystéine. Cette activité protéolytique est nécessaire pour la reconnaissance et l'adhésion du parasite à la surface des cellules de l'hôte. Mendoza Lopez et al (Mendoza-Lopez et al., 2000) ont mis en évidence une protéase à cystéine de 30 kDa (CP30) possédant une forte affinité avec la surface des cellules HeLa. L'utilisation d'un anticorps anti-CP30 entraîne une diminution de l'adhésion du parasite pouvant aller jusqu'à 50%. De même, traiter des parasites avec des inhibiteurs spécifiques des protéases à cystéine comme l'E64 diminue la cytoadhérence du parasite. Des résultats similaires ont été obtenus par l'équipe d'Hernandez (Hernandez et al., 2004) avec une protéase de 62 kDa. L'injection à des souris d'un anticorps dirigé contre cette protéine, 24h avant l'infection par les parasites, entraînait une protection de 92% des souris.

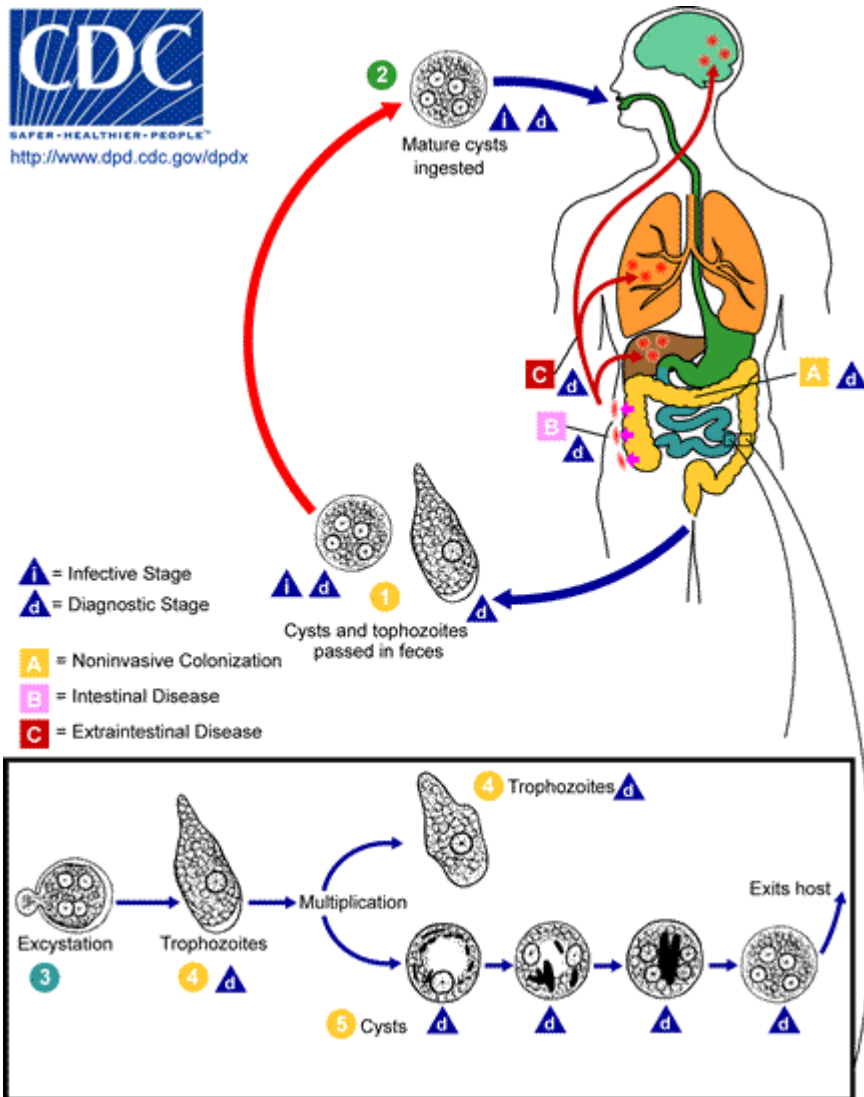


Figure 19 : cycle de développement d'*Entamoeba histolytica*. Les kystes et les trophozoïtes sont retrouvés dans les fèces (1). L'infection par *Entamoeba histolytica* se produit par l'ingestion de kystes mûrs (2). L'excystation (3) se produit dans le petit intestin, libérant des trophozoïtes qui migrent au niveau du gros intestin (4). Les trophozoïtes se multiplient par fission binaire et produisent des kystes (5). Les kystes peuvent survivre plusieurs jours dans l'environnement externe et sont responsables de la transmission. Les trophozoïtes passés dans les fèces sont rapidement détruits dans l'environnement extérieur, et s'ils étaient ingérés, ne survivraient pas à l'environnement gastrique. Dans beaucoup de cas, les trophozoïtes restent confinés au lumen intestinal (A : infection non envahissante). Dans quelques cas les trophozoïtes envahissent la muqueuse intestinale (B: maladie intestinale), ou, au moyen de la circulation sanguine, ils vont envahir d'autres organes tels que le foie, le cerveau, et les poumons (C: maladie extra-intestinal), et provoquer des manifestations pathologiques.

2.1.2 Implication des protéases dans l'invasion tissulaire et cellulaire

Les parasites, pour pouvoir perdurer dans un organisme, migrer, échapper à la réponse immunitaire ou en raison d'un stade intracellulaire lors de leur cycle de développement, emploient des mécanismes sophistiqués d'invasion des tissus ou des cellules.

Par, exemple, l'événement initial de l'adhésion d'*E. histolytica* sur une cellule cible, correspond à l'interaction de glycoprotéines présentes à la surface des trophozoïtes (Figure 19) avec des sucres présents à la surface de la cellule (Variyam, 2007). Le mucus du côlon humain possède des sucres similaires aux cellules. L'attachement des trophozoïtes au mucus peut prévenir leur attachement aux cellules épithéliales. En réponse à ce mécanisme, *E. histolytica* va sécréter des protéases capables de dégrader ce mucus et d'altérer sa fonction protectrice. Ainsi, les protéases à cystéine d'*E. histolytica* sont capables de dégrader les régions faiblement glycosylées des mucines. La mucine ainsi dégradée perd son efficacité à inhiber l'adhésion des amibes aux cellules (Moncada et al., 2003).

Un autre exemple concerne le rôle des protéases chez les plasmodies de *Plasmodium*. L'invasion des globules rouges par ce pathogène est une étape essentielle du cycle de développement de ce parasite intracellulaire obligatoire (Figure 20). Un grand nombre de protéases à sérine et cystéine sont impliquées dans les processus d'invasion de cet organisme. L'étude des effets de différents inhibiteurs de protéases suggère un rôle de celles-ci dans l'invasion et la rupture des érythrocytes par le stade mérozoïte. L'utilisation d'inhibiteurs de protéases à sérine, tel que la chymostatine, ou de protéases à cystéine, notamment les inhibiteurs spécifiques de la falcipaine, sont capables d'inhiber l'invasion des érythrocytes par les plasmodies (Greenbaum et al., 2002; Rosenthal, 2004). Une protéase à sérine, PfsUB1 joue un rôle important dans cette phase d'invasion des érythrocytes (Koussis et al., 2009). Au stade érythrocytaire, les parasites se répliquent à l'intérieur d'une vacuole parasitophore. La rupture de la membrane de la cellule-hôte va libérer des mérozoïtes-filles, capables d'infecter de

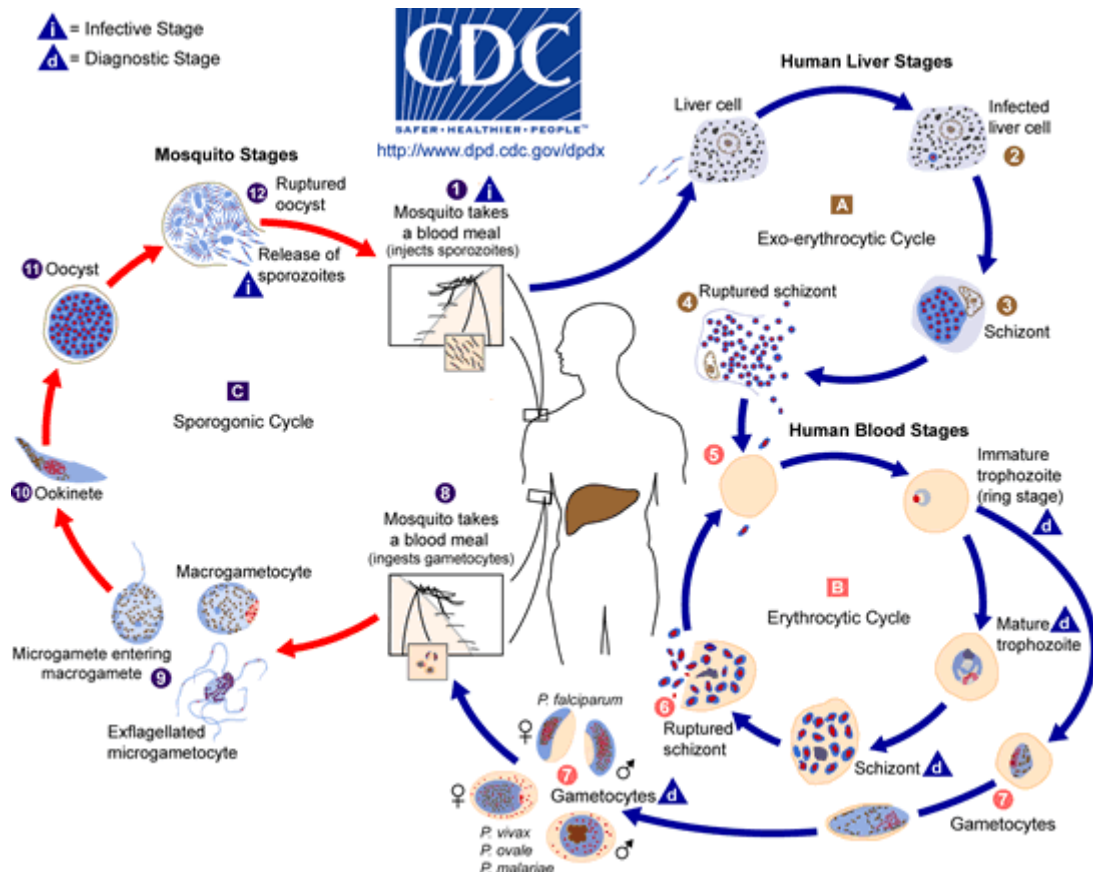


Figure 20 : cycle de développement des parasites du genre *Plasmodium*. Le cycle de vie du parasite implique deux hôtes. Pendant un repas de sang, un moustique femelle infecté inocule des sporozoïtes dans l'hôte humain (1). Les sporozoïtes infectent des cellules hépatique (2) et mûrissent en schizontes (3), qui une fois rompus vont libérer des mérozoïtes (4). Après cette réplication initiale extra érythrocytaire (A), les parasites entre dans un cycle de multiplication asexué dans les érythrocytes (schizogonie érythrocytaire (B)). Les mérozoïtes (5) infectent les globules rouges. Les trophozoïtes mûrissent en schizontes, dont la rupture va libérer des mérozoïtes (6). Quelques trophozoïtes immatures vont se différencier (gamétocytes) (7) pour entrer dans un cycle de multiplication sexuée dans les érythrocytes. Les stades sanguins du parasite sont responsables des manifestations cliniques de la maladie. Les gamétocytes, mâles (micro-gamétocytes) et femelles (macro-gamétocytes), sont ingérés par un moustique pendant un repas de sang (8). La multiplication des parasites dans le moustique est connue comme cycle sporogonique (C). Dans l'estomac du moustique les micro-gamètes fécondent les macro-gamètes produisant des zygotes (9). Les zygotes deviennent motiles et allongés (ookinètes)(10) puis se développent en ookystes (11) . Les ookystes croissent, puis rompent libérant des sporozoïtes (12), qui vont migrer jusqu'au glandes salivaires du moustique. L'inoculation des sporozoïtes dans un nouvel hôte humain perpétue le cycle de vie du parasite.

nouveaux érythrocytes. Pour initier cette invasion, les mérozoïtes expriment à leur surface trois protéines MSP1, 6, et 7 (merozoite surface proteins), impliquées dans les interactions avec les érythrocytes. Juste avant la libération des mérozoïtes-filles, ces protéines vont subir une maturation protéolytique par PfSUB1. L'inhibition de PfSUB1 entraîne une accumulation de protéines MSPs non effectives à la surface des mérozoïtes et une diminution significative de l'invasion des érythrocytes. Cette protéine joue donc un rôle essentiel dans l'invasion cellulaire à travers le remodelage de protéines de surface.

2.1.3 Le dékystement

La formation de kystes est une des stratégies de survie des parasites à l'extérieur de leur hôte. Il s'agit de la forme de résistance et de contamination dans les cas de transmission par voie orale. Pour pouvoir débiter son cycle le parasite doit pouvoir émerger du kyste. Pour bon nombre de protozoaires, helminthes ou arthropodes, les protéases sont nécessaires pour ce dékystement.

L'amibe *Giardia lamblia*, agent responsable de la giardiose chez l'homme, initie son cycle par la rupture de la paroi kystique sous l'action de protéases à cystéine (Sajid et McKerrow, 2002). Des protéases à cystéine ont également été identifiées dans les œufs de *Schistosoma mansoni*, ver parasite de l'homme responsable de la bilharziose, et semblent impliquées dans l'éclosion des œufs et l'invasion des tissus de l'hôte (Sajid et McKerrow, 2002). L'utilisation d'inhibiteurs de protéases (DTT, N-éthylmaléimide, PMSF, cystatine) sur une autre famille de parasite, les cryptosporidies, et notamment sur les ookystes de *Cryptosporidium parvum*, entraîne une inhibition du dékystement *in vitro*, ainsi que de l'infestation (Kniel et al., 2004).

2.1.4 Les protéases comme facteurs de virulence

2.1.4.1 *Virulence et facteurs de virulence : définitions*

La virulence est un facteur difficile à définir. Le mot « virulence » est couramment utilisé pour caractériser la capacité relative d'un organisme à provoquer une maladie et a été traditionnellement employé pour décrire une caractéristique microbienne. Cette définition de la virulence est basée sur des caractéristiques du pathogène, incluant son degré de pathogénie, sa capacité à surmonter le système immunitaire, son pouvoir invasif, la sévérité des maladies qu'il induit, les dommages qu'il engendre, ou encore le pourcentage de mortalité induite par l'infection. Cette capacité du pathogène à causer une maladie a ainsi été attribuée à l'expression de facteurs dits de virulence. Ces facteurs ont été définis classiquement comme des composants d'un pathogène qui une fois supprimés altèrent sa virulence mais pas sa viabilité (Poulin et Combes, 1999). Pourtant, il existe une certaine confusion liée à cette définition, confusion qui provient justement du fait que la virulence est considérée comme une propriété microbienne, alors qu'elle semble exprimée seulement chez des hôtes susceptibles. De ce point de vue, le concept qui définit la virulence comme une propriété microbienne intrinsèque distinguant les organismes pathogènes des non pathogènes devient difficile à appliquer, notamment face au nombre croissant de preuves montrant les facteurs de l'hôte comme des déterminants critiques de la relation hôte pathogène. Ainsi, un organisme dit virulent peut ne pas l'être chez un hôte possédant une immunité spécifique, et un organisme non virulent peut provoquer une infection chez un hôte avec une immunité altérée (Casadevall et Pirofski, 2001). Par conséquent, plutôt que de définir la virulence comme une variable indépendante, nous la définirons comme une variable dépendante de la susceptibilité de l'hôte, ainsi que du contexte et de la nature de la relation hôte-pathogène.

2.1.4.2 Les protéases facteurs de virulence

Chez les différentes espèces d'*Entamoeba* spp., les protéases à cystéine sont considérées comme des facteurs importants de la virulence. Cette notion est appuyée par les études montrant une activité protéasique de type protéase à cystéine de 10 à 1000 fois supérieure dans les lysats parasitaires de l'amibe pathogène *E. histolytica* par rapport aux lysats de l'amibe non pathogène *E. dispar*. Ces protéases à cystéine sont capables de dégrader le collagène, le fibrinogène, les éléments de la matrice extracellulaire mais également d'avoir un effet cytopathique sur cellules *in vitro*. Cet effet sur des fibroblastes en culture est complètement inhibé par du benzyloxycarbonyl-Phe-Arg-fluoromethylketone (Z-Phe-Arg-CH₂F), inhibiteur spécifique des protéases à cystéine, non toxique pour les cellules de l'hôte (Que et Reed, 2000). Le rôle des cystéines à protéases comme facteurs de virulence a également été montré *in vivo*. Stanley's et al. (Stanley et al., 1995) ont ainsi montré que l'inhibition de l'activité des protéases à cystéine diminuait la formation d'abcès hépatique chez des souris immunodéficientes. Sur un autre modèle animal, le hamster, il a été montré que l'injection de trophozoïtes dans la veine porte d'animaux recevant une dose constante d'E64, protégeait les tissus d'éventuelles lésions, et entraînait une rapide élimination des parasites (Olivos-Garcia et al., 2004). Ces travaux suggèrent que l'activité des protéases à cystéine est nécessaire pour la survie des parasites dans des conditions *in vivo*, et pour les phénomènes de virulence.

Chez les trichomonadines, les protéases à cystéine sont également impliquées dans plusieurs aspects de leur virulence. Chez *Trichomonas vaginalis*, il est maintenant avéré que les protéases à cystéine CP65 et CP30 participent à l'effet cytotoxique des parasites sur des cellules HeLa ou des cellules de l'épithélium vaginal, selon un mode dépendant du contact hôte/parasite (Arroyo et Alderete, 1995). L'utilisation d'un anticorps anti-CP65 en prétraitement sur des parasites, réduit leur cytotoxicité sur les cellules HeLa (Alvarez-Sanchez et al., 2000). Il est de même avec l'utilisation d'E64, inhibiteur spécifique des protéases à cystéine, confirmant ainsi que la protéine CP65 est

Tableau 6 : Mécanismes d'évasion au système immunitaire employés par certains protozoaires parasites (Zambro-Villa et al, 2002).

Parasite (disease)	Epidemiology of the disease [1]	Main strategies of evasion	Result	Refs
<i>Plasmodium falciparum</i> (malaria)	300 million–500 million infected per year 1.5 million–2.7 million deaths per year	Antigenic variation and/or polymorphisms	Evades the IR; infected cells are prevented from being swept to the spleen	[2]
		Induction of blocking antibodies	Blocks the binding of real inhibitory antibodies	[4]
		Molecular mimicry	Alters immune recognition and might induce autoimmune disease	[6,7]
		Anergy of T cells Altered peptide ligand	Immunosuppression Alters functions of memory T cells	[10–14] [15,16]
<i>Trypanosoma brucei</i> (African trypanosomiasis or sleeping sickness)	15 000–20 000 new cases per year 55 million at risk	Antigenic variation by VSG	Evades the previously established IR	[17,18]
		Alteration of T- and B-cell populations	Immunosuppression	[19]
		Abnormal activation of macrophages	Impairs macrophage functions	[20]
		Induction of changes in the pattern of cytokines released by CD8 ⁺ T cells Production of a gp63-like-protein	Increases IFN- γ and decreases IL-2 and IL-2R; renders T cells unresponsive Resists complement	[21] [22]
<i>Trypanosoma cruzi</i> (Chagas disease)	12 million–16 million infected 45 000 deaths per year 90 million at risk	Increased phagocytic activity	More CD8 ⁺ T cells and reduced TDR and TIR	[23]
		Parasite mucin that binds to macrophage	Impairs macrophage cytokine secretion	[25]
		Anergy of T cells Production of blocking IgM antibodies	Immunosuppression Blocks the binding of real inhibitory antibodies	[26,27] [29]
		Turnover of surface molecules, phospholipases and complement-regulating factors	Resists complement	[31–33]
<i>Entamoeba histolytica</i> (Intestinal and liver amoebiasis)	500 million infected 40 000–110 000 deaths per year	Cytolytic capacity	Damages host cells and tissues, interfering with IR	
		Degradation of antibodies by proteases	Evades humoral immunity	[34]
		Acquisition of complement-regulating factors, shedding of immune complexes by capping and inactivation of C3a and C5a	Resists complement and protects against the inflammatory response	[36,37]
		Anergy of T cells	Immunosuppression	[38,39] [41–43]
		Release of products (MLIF among others) that act on macrophages; produces PGE2	Impairs macrophage functions	
		Induction of IL-4 and IL-10	Modulates the Th1 response	[45]
<i>Leishmania</i> parasites (leishmaniasis: cutaneous, mucocutaneous and visceral)	12 million infected 350 million at risk 1 million–1.5 million new cases per year of cutaneous leishmaniasis 500 000 new cases and 75 000–80 000 deaths per year owing to visceral leishmaniasis	Inhibition of phagolysosome formation and the proteolytic enzymes from lysosome	Evades the macrophage proteolytic processes	[46,47]
				[48]
		Abnormal activation of protein kinase C and scavenging of oxygen intermediates	Inhibits the respiratory burst	[49]
		Prevention of apoptosis of infected macrophage	Extends the survival of the infected macrophage	[50,51,53]
		Inhibition of MHC-protein production, peptide loading and expression of co-stimulatory molecules	Impairs macrophage antigen-presenting function	[54,56] [57]
				[58,59]
		Induction of PGE2 and downregulation of TNF- α R	Impairs macrophage function	
		Inhibition of macrophage and neutrophil chemotaxis	Promotes survival of parasites once established	[61]
		Shedding of MAC complex and inactivation of some MAC components Suppression of transcription of the <i>IL-12</i> gene	Resists complement Blocks the protective Th1 response	

*Abbreviations: IFN- γ , interferon γ ; IL-2, interleukin 2; IL-4, interleukin 4; IL-10, interleukin 10; IL-12, interleukin 12; IL-2R, IL-2 receptor; IR, immune response; MAC, membrane attack complex; MHC, major histocompatibility complex; MLIF, monocyte locomotion inhibition factor; PGE2, prostaglandin E2; TDR, thymus-dependent response; TIR, thymus-independent response; TNF- α R, tumor necrosis factor α receptor; VSG, variant surface glycoprotein.

une protéase à cystéine. Cette protéase qui possède une activité optimale à pH 5,5 et à 37°C, conditions similaires à celles retrouvées chez les patientes présentant une trichomonase, est capable de dégrader quelques unes des protéines constituant l'épithélium vaginal, le collagène IV et la fibronectine.

2.2 Comment les parasites échappent à la réponse immune

Le succès de la survie de nombreux protozoaires pathogènes comme *Plasmodium* spp., *Leishmania* spp., *Trypanosoma* spp., *Entamoeba* spp., *Giardia* spp., ou encore *Trichomonas* dépend pour beaucoup de leur capacité à échapper à la réponse immune de l'hôte. Certains mécanismes employés mettent en jeu des protéases, notamment dans le cas de processus d'immunosuppression et de modulation de la réponse immune (Zambrano-Villa et al., 2002) (Tableau 6).

2.2.1 Immunosuppression

L'importance des protéases dans l'évasion du système immunitaire de l'hôte a été reconnue pour la première fois par Chapman et Mitchell (Chapman et Mitchell, 1982), après observation *in vitro* du clivage des chaînes lourdes d'immunoglobulines G murines par des produits de sécrétions de *Fasciola hepatica*. Depuis, nombre d'exemples de clivage *in vitro* des immunoglobulines par des protéases ont été documentés, chez les parasites *Trichomonas* spp., *Entamoeba* spp., *Giardia* spp., *Trypanosoma* spp., *Fasciola* spp., ou *Taenia* spp. Dès lors l'hypothèse admise est que la sécrétion *in vivo* de ces protéases par les parasites jouait un rôle de protection face à la réponse immunitaire. Les parasites qui ont besoin d'adhérer aux muqueuses pour pouvoir envahir les tissus et perdurer dans l'intestin ont par exemple développés des mécanismes de clivage des immunoglobulines A. Celles-ci sont souvent la première ligne de défense des muqueuses contre les agents pathogènes invasifs (Woof et Kerr, 2006). Elles sont la principale classe d'anticorps présents dans les sécrétions mucosales, et possèdent une grande variété de fonctions telles que diminuer la colonisation des muqueuses, neutraliser les toxines

produites, ou diminuer l'activité protéolytique des enzymes produites par les agents pathogènes. Pourtant, la fonctionnalité de ces immunoglobulines A dépend de leur intégrité. Les parasites ont donc développé des mécanismes de destruction de celles-ci. Les trophozoïtes d'*E. histolytica*, par exemple, sécrètent à la surface de leurs membranes des protéases à cystéine susceptibles de dégrader les IgA secrétées humaines (Garcia-Nieto et al., 2008). Des protéases retrouvées dans le lysat et le surnageant de culture de *T. vaginalis* possèdent la capacité de dégrader les chaînes lourdes des IgA (Klemba et Goldberg, 2002). Les autres immunoglobulines sont également la cible des protéases. En effet, bon nombre de parasites relâchent des protéases capables de cliver les IgG, les rendant ainsi inactives. *T. cruzi* est capable de se lier au fragment Fab des IgG et de cliver ensuite le fragment Fc au moyen de la cystéine protéase cruzipaine (Bontempi et Cazzulo, 1990). Le fragment Fab des immunoglobulines est le fragment fixant l'antigène. Le fragment Fc possède lui des fonctions multiples comme l'activation du complément ou l'opsonisation⁶. Le clivage des immunoglobulines peut également être illustré par le ver parasite *Schistosoma mansoni* (Aslam et al., 2008). Les infections à *Schistosoma* sont souvent associées avec un titre anticorps élevé en IgE. Un lien direct a été établi chez l'homme entre un taux élevé d'IgE et l'acquisition d'une résistance face à une réinfection par des schistosomes, de même que l'implication des IgE dans la mort du parasite. Pourtant malgré une réponse immune active, les schistosomes sont capables de survivre une longue période dans la circulation sanguine, laissant supposer que le parasite est capable d'échapper à la réponse immunitaire. Un des mécanismes utilisés par le parasite est le clivage des IgE en fragments non fonctionnels par des enzymes. L'utilisation de substrat peptidique synthétique et d'inhibiteurs suggèrent l'implication d'une protéase à sérine de type chymotrypsine (Pleass et al., 2000).

Un autre rôle des protéases se situe au niveau de l'évasion à la lyse par le complément. Les nématodes et les trypanosomatidae sécrètent des protéases capables

⁶ processus biologique favorisant la phagocytose. Les Ig reconnaissent des épitopes situés à la surface d'un pathogène, par leur fragment Fab. Les récepteurs des phagocytes vont alors reconnaître le fragment Fc des Ig, permettant ainsi la phagocytose

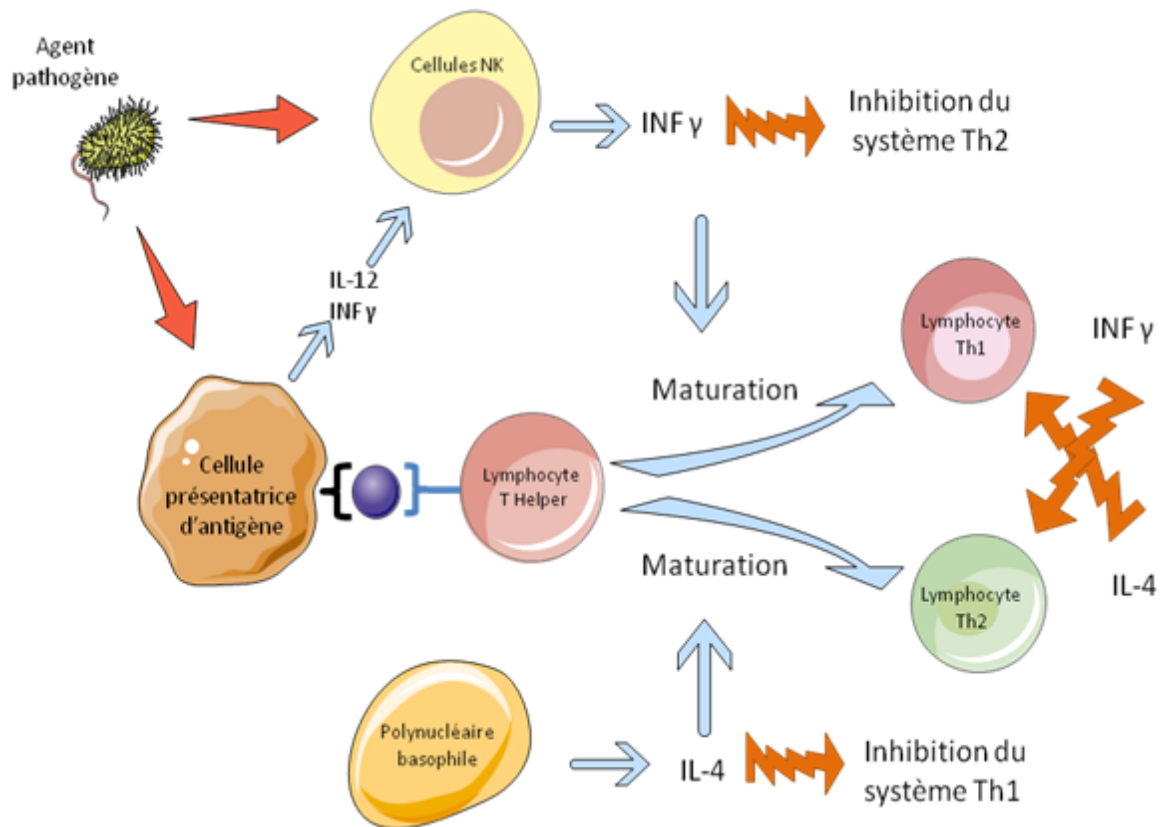


Figure 21 : Schéma simplifié de la différenciation des lymphocytes T précurseurs.

Selon l'environnement dans lequel ils vont se trouver, les lymphocytes Th vont soit se différencier en lymphocytes Th1 soit en lymphocytes Th2. Les lymphocytes Th1 orientent la réponse vers l'immunité à médiation cellulaire alors que les Th2 orientent la réponse vers l'immunité à médiation humorale. Il existe un équilibre entre les 2 réponses, puisque la production d'IL-4 par les Th2 va bloquer la réponse médiée par les Th1, et que l'INF γ produit par les Th1 va bloquer la réponse médiée par les Th2.

d'inactiver des protéines composant le complément, se protégeant ainsi d'une lyse par celui-ci. *Leishmania* spp, par exemple, possèdent des métalloprotéases dénommées MSP (major surface protease) ayant une forte affinité de liaison avec la protéine C3 du complément. L'activité protéolytique de ces protéases va entraîner une conversion de la forme C3b active vers une forme inactive protégeant ainsi le parasite de toute lyse par le complément (Yao, 2010). Un système similaire est utilisé par les nématodes *Brugia malayi* et *Trichinella spiralis* (Rees-Roberts et al., 2010) dont les produits sécrétés par les stades larvaires vont influencer le recrutement des granulocytes par la protéine C5a du complément. Ce recrutement est même complètement aboli dans le cas de *Brugia malayi*. Cette inhibition du recrutement est totalement bloquée par l'utilisation de PMSF, un inhibiteur spécifique des protéases à sérine, indiquant l'implication de ces dernières, qui agissent par clivage de la protéine C5a. Dans le cas de *Trichinella spiralis*, l'inhibition du recrutement est plus modeste et impliquerait des métalloprotéases.

2.2.2 Immunomodulation de la réponse immune

La modulation des mécanismes de défense de l'hôte par les organismes pathogènes est reconnue depuis plusieurs années. Les analyses de la réponse immunitaire face à ces organismes révèlent un rôle important des cytokines et chimiokines en temps que régulateurs de la réponse inflammatoire. Les protéases sécrétées par les parasites sont capables de moduler la production de ces molécules de manière à pouvoir en tirer un bénéfice (Mottram et al., 1998). Cette modulation peut se faire aussi bien par stimulation de l'expression de cytokines que par le clivage de celles-ci.

Ainsi, les protéases à cystéine des parasites inhibent les lymphocytes T auxiliaires Th1 par l'induction d'interleukine 4. En effet, l'IL-4 sécrétée par des granulocytes basophiles et par les lymphocytes T auxiliaires Th2 est un inhibiteur du système Th1 (Choi et Reiser, 1998) (Figure 21). Les lymphocytes Th1 sont connus pour orienter la réponse immunitaire vers l'immunité à médiation cellulaire, et vont activer les

macrophages qui sont responsables de la neutralisation des micro-organismes pathogènes. Les protéases parasitaires qui agissent en activant les granulocytes basophiles provoquent leur dégranulation et la sécrétion d'IL-4 malgré l'absence de présentation d'antigènes spécifiques (Dixit et al., 2008). En favorisant la sécrétion d'IL-4, les parasites se protègent de l'immunité à médiation cellulaire. Les protéases à cystéine interviennent également dans la réponse immunitaire médiée par les lymphocytes auxiliaires Th2 (immunité à médiation humorale). Chez *Leshmania mexicana*, on observe rapidement lors de l'infection de souris, une production d'IL-4 et le développement de la réponse immune médiée par les Th2. L'IL-4 produite par les Th2 bloquent la réaction immunitaire médiée par les Th1. Le blocage de l'expression du gène de l'IL-4 inhibe la progression de l'infection (Satoskar et al., 1995). De plus, des mutants de *L. mexicana* délétés des gènes codant les protéases à cystéine, induisent une réponse immunitaire médiée par les Th1 chez les souris. Ceci se traduit par un pouvoir infectieux plus faible et l'absence de lésions (Alexander et al., 1998).

Une autre voie utilisée par les parasites pour moduler la réponse immunitaire est le clivage des cytokines/chimiokines par des protéases. Ce mécanisme a été largement étudié chez *E. histolytica*. L'effet de la protéase à cystéine EhCP5 spécifique à *E. histolytica* sur l'IL-18 a notamment été étudié (Que et al., 2003). L'IL-18 est connue pour posséder un rôle central en stimulant des cellules NK et les lymphocytes T, qui vont alors produire de l'interféron gamma et ainsi activer les macrophages (Dinarello, 1999). L'EhCP5 est capable de cliver la proIL-18 et l'IL-18 mature humaine, les rendant ainsi biologiquement inactives. De même la protéase EhCP2 est capable de cliver les chimiokines CCL2, CCL13 et CXCL8 (Pertuz Belloso et al., 2004). Les chimiokines sont caractérisées par leurs propriétés chimio-attractantes (Luster, 1998). Les CCL2 activent les monocytes et lymphocytes qui expriment le récepteur CCR2 de ces chimiokines. Les CCL2 et les CCL13 sont également impliquées dans le développement des maladies inflammatoires (Garcia-Zepeda et al., 1996). Les CXCL8 activent les granulocytes et monocytes qui expriment CXCR1 (Pertuz Belloso et al., 2004). Ces clivages par EhCP2

mènent à une diminution de l'activité biologique des différentes chimiokines, entraînant par la même une diminution du recrutement et de l'activation des leucocytes. Ce mécanisme peut être vu comme un mécanisme de régulation utilisé par *Entamoeba* pour éviter la réponse immune de l'hôte.

3. Les protéases, des cibles intéressantes dans la lutte contre les parasites ?

Essentielles pour le fonctionnement des parasites, les protéases sont très vite devenues des cibles de choix dans l'élaboration de traitements antiparasitaires et de diagnostic.

3.1 Des cibles thérapeutiques

Nous avons pu voir précédemment que les protéases étaient un groupe conséquent, varié, et ubiquitaire d'enzymes jouant des rôles-clés dans la plupart des processus biologiques des protozoaires parasites. En tant que cibles potentielles de molécules thérapeutiques, elles présentent de nombreux avantages (Sajid et McKerrow, 2002) : leurs mécanismes catalytiques, leurs fonctions, la spécificité des substrats sont autant d'informations connues depuis des années. De plus, des protéases appartenant à une même famille présentent souvent des homologues chez les différents parasites. Il paraît alors raisonnable de penser que des inhibiteurs puissent être effectifs chez différents parasites. Certaines protéases parasitaires présentent des homologues chez les mammifères qui sont utilisés comme cibles thérapeutiques par les laboratoires pharmaceutiques, pour leur implication dans les pathologies humaines. Des inhibiteurs de protéases sont utilisés pour le traitement de l'hypertension (McKerrow et al., 2008). Des essais cliniques sont en cours pour la validation de protéases comme cibles pour traiter le diabète, l'ostéoporose ou les cancers. Toutes les informations disponibles offrent donc la possibilité d'approches très rentables pour le développement de nouveaux antiparasitaires. Il convient aussi de travailler sur la sélectivité des

Tableau 7 : Exemples de quelques protéases ciblées chez des protozoaires parasites pathogènes.

Parasite	Protease	Discovery	Hits	Leads	Validation *	References
<i>Trypanosoma cruzi</i>	Cruzain	X	X	X	X	[12,16,17,25,26••,33]
	Cathepsin B	X	X			
	OPB	X	X			
	Carboxypeptidas [fix break]	X				
<i>Trypanosoma brucei</i>	Brucipain	X	X	X	X	[27]
	TbCatB	X	X	X	X	
Leishmania	CpB	X	X		X	[10,11,22,24]
	OPB	X	X			
	Subtilisin	X				
Toxoplasma	X		X	X		[5]
	Falcipains	X	X	X	X	
<i>Plasmodium falciparum</i>	Plasmepsins	X	X			[4•,14]
	Falcilysin	X	X			
<i>Entamoeba histolytica</i>	Cysteine protease	X	X			[21]

* La cruzipaine chez *T. cruzi*, la brucipaine et TbCatB chez *T. brucei*, La CpB chez *Leishmania*, la facipaine chez *P. falciparum* se sont avérées des cibles intéressantes pour des inhibiteurs (McKerrow et al., 2008).

composants développés envers les cibles parasitaires (Zucca et Savoia, 2011). Les stratégies visant à inhiber spécifiquement les protéases parasitaires ont cependant déjà fait leurs preuves (Tableau 7). De plus la disponibilité de séquences de génomes complets des parasites et de leurs hôtes offre à ces stratégies de nouvelles promesses pour le développement de nouveaux antiparasitaires, en permettant la mise en évidence des enzymes spécifiques aux parasites et essentielles à leur biologie.

3.2 Des cibles vaccinales

Contrôler une maladie par vaccination en lieu et place de l'utilisation de molécules antiparasitaires est une solution moins coûteuse, plus efficace et plus fiable à long terme pour la prévention d'une infection et de sa transmission (Rogier et al., 2007). Un autre avantage à la vaccination est son mode d'administration, bien moins contraignant que la mise à disposition immédiate d'une chimio-prophylaxie. De plus, la propagation des résistances aux molécules antiparasitaires a rendu nécessaire le développement de nouvelles solutions.

Le rôle important des protéases, la possibilité d'inhiber l'activité protéolytique de celles-ci au moyen d'anticorps spécifiques et leurs propriétés antigéniques en ont fait des cibles de choix pour les vaccins. Ainsi, chez le mouton, l'utilisation d'une cathepsine L recombinante comme antigène pour un vaccin dirigé contre *Fasciola hepatica*, avec l'administration à trois reprises de 10µg de cathepsine L induit une réponse protectrice (Dixit et al., 2008). Cette vaccination permet une diminution de l'ordre de 50% de la charge parasitaire en comparaison des témoins non vaccinés (Golden et al., 2010). Chez *E. histolytica*, des essais vaccinaux chez les rongeurs ont été fait avec la protéase EhCP112, connu pour former un complexe avec une adhésine à la surface des amibes (Lotter et Tannich, 2006). Une immunisation avec un plasmide codant la protéase et l'adhésine provoque une forte réponse IgG. La vaccination intra-hépatique chez le hamster montre une réduction significative des abcès hépatiques. Chez les leishmanies, parasites où le développement de vaccins a été marqué par des progrès notables, des

vaccins constitués d'antigènes parfaitement définis sont en cours de développement. Parmi ces antigènes l'utilisation de la métalloprotéase GP63 s'avère prometteuse, et son utilisation en vaccination confère une immunité à des souris BALB/c face à *L. donovani* (Kaur et al., 2011).

Pourtant, si la vaccination antiparasitaire est une alternative intéressante à la chimio-prophylaxie, aucun vaccin n'est à l'heure actuelle disponible chez l'homme. La difficulté de mettre au point ces vaccins et le peu d'intérêt que porte l'industrie pharmaceutique en raison d'un marché peu solvable, ont fait que la recherche n'en ait qu'au stade des essais cliniques (Rogier et al., 2007).

3.3 Des marqueurs pour le séro-diagnostic

Détecter rapidement une infection est crucial pour le contrôle d'une parasitose. En effet poser le diagnostic oriente la prise en charge thérapeutique. La méthode standard de détection d'une infection causée par des protozoaires ou des helminthes est la microscopie. Pourtant, utiliser une méthode immunologique peut révéler d'autres informations, comme le titre en anticorps, le niveau de l'infection, ou l'évolution de celle-ci, se révélant ainsi utile pour l'adaptation d'un traitement.

Les protéases ayant la caractéristique d'être très immunogènes, elles furent rapidement exploitées comme marqueur sérologique d'infection pour la mise au point de tests de diagnostic. Les protéases à cystéine ont notamment été testées comme marqueurs sérologiques pour la mise en évidence des infections à *F. hepatica*, *F. gigantica*, *E. histolytica*, *T. vaginalis*, *L. major*, *T. Cruzi* ou encore à *P. falciparum* (Sajid et McKerrow, 2002).

Les protéases parasitaires, en plus d'être d'une grande diversité dans leurs modes de fonctionnements sont des enzymes ayant de multiples rôles biologiques. Les quelques exemples développés ici, ne sont qu'une liste non exhaustive de leur

implication dans la pathogénie des organismes parasites. L'intérêt croissant que portent les chercheurs à ces protéases, notamment à des fins thérapeutiques, est la meilleure illustration de leur rôle fondamental chez les organismes pathogènes. Dès lors, il nous a semblé intéressant d'orienter nos recherches sur les protéases de *Blastocystis* sp. ST7, afin d'apporter des arguments sur son caractère pathogénique. Le séquençage complet du génome de *Blastocystis* sp. ST7 nous a permis d'obtenir de précieuses informations pour l'identification de ces facteurs de virulence potentiels.

Chapitre II.

Résultats expérimentaux

I. Séquençage des génomes nucléaire et mitochondrial de *Blastocystis* sp. sous-type 7.

Les données génomiques d'un organisme constituent une source importante d'informations pour la compréhension des potentialités de celui-ci. Le séquençage du génome de *Blastocystis* sp. ST7 avait ainsi pour objectif d'apporter des éléments de réponse concernant son caractère potentiellement pathogène et de décrypter certains métabolismes particuliers, comme ceux du MLO, qui pourraient servir de cibles thérapeutiques.

Ce projet, réalisé en collaboration avec le Génoscope d'Evry, l'Institut Pasteur de Lille, l'Université Nationale de Singapour et l'Université de Provence, a nécessité au préalable deux années de travail pour la mise en place de la culture au laboratoire et pour la production de l'ADN nécessaire au séquençage du parasite. L'isolat de *Blastocystis* sp. provient des selles d'un patient symptomatique singapourien, et correspond au sous-type 7 selon la terminologie établie par Stensvold et al. (Stensvold et al., 2007b).

La résolution du caryotype de *Blastocystis* sp. ST7 par électrophorèse en champ pulsé a montré 15 bandes chromosomiques. L'assemblage des séquences a abouti à 54 unités (scaffolds), pour une taille totale de 18,8 Mb (couverture de 16,5 x). Le détail descriptif de ce génome a fait l'objet d'une publication dans *Genome Biology*, dont je suis signataire (Denoëud et al., 2011). L'étude de ce génome a permis d'apporter des pistes pour comprendre l'adaptation de ce parasite, qui est le seul straménopile parasite de l'homme. Des génomes de straménopiles libres ou parasites de végétaux étant disponibles, l'étude de ce génome est intéressante pour comprendre l'évolution génomique au sein d'un groupe dont les membres ont eu des trajectoires évolutives très différentes et comprendre les mécanismes requis pour qu'un microorganisme puisse survivre et persister au niveau intestinal. Sur le plan évolutif, les conséquences pour le

parasite sont très souvent une régression génique induisant une régression de certaines fonctions métaboliques.

L'analyse complète du génome de *Blastocystis* sp. ST7 a révélé plusieurs mécanismes d'évolution suggérant une adaptation du parasite à son environnement. Ainsi, ce génome comporte une large proportion de gènes dupliqués, probablement issus d'une duplication complète du génome ou de duplications segmentaires, suivies de réarrangements. Un autre mécanisme, important en termes d'évolution génomique, mis en évidence chez *Blastocystis* sp. ST7, est le transfert horizontal de gènes. Ce mécanisme joue un rôle important en termes d'acquisition de nouveaux gènes et donc de nouvelles potentialités. L'analyse des données a permis d'identifier 133 gènes candidats présentant exclusivement des homologues de séquences avec des gènes bactériens et d'archaea. Il semble également que *Blastocystis* sp. ST7 ait acquis des gènes par transfert entre eucaryotes.

Ce modèle parasitaire est également intéressant d'un point de vue de l'évolution des organites apparentés aux mitochondries, du fait de la présence chez un organisme anaérobie de MLOs possédant des caractéristiques morphologiques proches des mitochondries et dont les métabolismes ne sont pas élucidés. Une molécule d'ADN circulaire de 29270 pb a été caractérisée au cours du séquençage et localisée dans les MLOs au moyen d'hybridations *in situ* en microscopie électronique à transmission. Cette molécule d'ADN réduite et compacte présente des caractéristiques mitochondriales : un pourcentage en AT de l'ordre de 80 %, une forte densité génique, des régions intergéniques très courtes et la présence de gènes codant des sous-unités du complexe I de la phosphorylation oxydative. En revanche, les gènes codant les autres complexes de la phosphorylation oxydative habituellement présents dans les mitochondries aérobies sont absents de ce génome mitochondrial (II, III, IV, ATP synthase). Cet aspect illustre la réduction génomique probablement liée aux contraintes environnementales, le parasite s'étant débarrassé des gènes non essentiels pour sa survie en anaérobiose. L'analyse de

ce génome est détaillée dans l'article publié dans la revue International Journal for Parasitology dont je suis 1^{er} auteur (Wawrzyniak et al., 2008).

La reconstruction du protéome *in silico* de ce MLO nous a permis d'éclairer les différents métabolismes se déroulant dans cet organite (Denoeud et al., 2011). Ainsi, nous avons identifié des protéines impliquées dans des fonctions typiquement mitochondriales (oxydation des acides gras, métabolisme des acides aminés, machinerie d'import mitochondrial, assemblage des protéines fer-souffre.....). En revanche l'absence des gènes codant les complexes III, IV et l'ATP synthase est confirmée par cette reconstruction, orientant vers la qualification de mitochondrie anaérobie pour cet organite. Par contre, les sous-unités du complexe II sont codées par le génome nucléaire et prédites pour être localisées dans le MLO. L'ensemble de ces données suggèrent que les MLOs de *Blastocystis* spp. ont été sélectionnés au cours de l'évolution en fonction des contraintes anoxiques de l'environnement intestinal.

RESEARCH

Open Access

Genome sequence of the stramenopile *Blastocystis*, a human anaerobic parasite

France Denoeud^{1†}, Michaël Roussel^{2,3†}, Benjamin Noel¹, Ivan Wawrzyniak^{2,3}, Corinne Da Silva¹, Marie Diogon^{2,3}, Eric Viscogliosi^{4,5,6,7}, Céline Brochier-Armanet^{8,9}, Arnaud Couloux¹, Julie Poulain¹, Béatrice Segurens¹, Véronique Anthouard¹, Catherine Texier^{2,3}, Nicolas Blot^{2,3}, Philippe Poirier^{2,3}, Geok Choo Ng¹⁰, Kevin SW Tan¹⁰, François Artiguenave¹, Olivier Jaillon¹, Jean-Marc Aury¹, Frédéric Delbac^{2,3}, Patrick Wincker^{1*}, Christian P Vivarès^{2,3*} and Hicham El Alaoui^{2,3*}

Abstract

Background: *Blastocystis* is a highly prevalent anaerobic eukaryotic parasite of humans and animals that is associated with various gastrointestinal and extraintestinal disorders. Epidemiological studies have identified different subtypes but no one subtype has been definitively correlated with disease.

Results: Here we report the 18.8 Mb genome sequence of a *Blastocystis* subtype 7 isolate, which is the smallest stramenopile genome sequenced to date. The genome is highly compact and contains intriguing rearrangements. Comparisons with other available stramenopile genomes (plant pathogenic oomycete and diatom genomes) revealed effector proteins potentially involved in the adaptation to the intestinal environment, which were likely acquired via horizontal gene transfer. Moreover, *Blastocystis* living in anaerobic conditions harbors mitochondria-like organelles. An incomplete oxidative phosphorylation chain, a partial Krebs cycle, amino acid and fatty acid metabolisms and an iron-sulfur cluster assembly are all predicted to occur in these organelles. Predicted secretory proteins possess putative activities that may alter host physiology, such as proteases, protease-inhibitors, immunophilins and glycosyltransferases. This parasite also possesses the enzymatic machinery to tolerate oxidative bursts resulting from its own metabolism or induced by the host immune system.

Conclusions: This study provides insights into the genome architecture of this unusual stramenopile. It also proposes candidate genes with which to study the physiopathology of this parasite and thus may lead to further investigations into *Blastocystis*-host interactions.

Background

Blastocystis sp. is one of the most frequent unicellular eukaryotes found in the intestinal tract of humans and various animals [1]. This anaerobic parasite was first described by Alexeieff at the beginning of the 20th century [2]. For a long time, the taxonomy of *Blastocystis* was controversial. Despite the application of molecular phylogenetic approaches, it was only recently that

Blastocystis sp. was unambiguously classified within the stramenopiles [3-5]. This eukaryotic major lineage, also called Heterokonta, encompasses very diverse organisms (unicellular or multicellular, heterotrophic or photosynthetic) such as slime nets, diatoms, water moulds and brown algae [6]. One important characteristic of stramenopiles is the presence during the life cycle of a stage with at least one flagellum permitting motility. It is important to note that *Blastocystis* sp. does not possess any flagellum and is the only stramenopile known to cause infections in humans [4]. For the organism isolated from human fecal material, Brumpt suggested the name *Blastocystis hominis* [7]. However, as the species *B. hominis* is difficult to establish, we use the term '*Blastocystis* sp.' to designate any subtype observed in

* Correspondence: pwincker@genoscope.cns.fr; christian.vivares@univ-bpclermont.fr; hicham.el_alaoui@univ-bpclermont.fr

† Contributed equally

¹Genoscope (CEA) and CNRS UMR 8030, Université d'Evry, 2 rue Gaston Crémieux, 91057 Evry, France

²Clermont Université, Université Blaise Pascal, Laboratoire Microorganismes: Génome et Environnement, BP 10448, F-63000 Clermont-Ferrand, France
Full list of author information is available at the end of the article

humans. *Blastocystis* sp. is the most frequent protozoa reported in human fecal samples [8], with a worldwide distribution [9-13] and a prevalence ranging between 30 and 60% in some developing countries [1]. In addition, infection with *Blastocystis* sp. appears to be common and more severe in immunocompromised or hemophilic patients [9,14,15]. The presence of *Blastocystis* representatives has also been reported in a variety of mammals, birds, reptiles, and even insects [16-18]. *Blastocystis* sp. exhibits extensive genetic diversity, and on the basis of molecular analysis of the small subunit RNA gene, ten distinct subtypes (ST1 to ST10) have been identified from primates (including humans), other mammals and birds [17]. Some arguments support zoonotic transmission to humans, including the high prevalence of ST1 to ST3 in humans and other mammals [17] and the experimental transmission of different human genotypes to chickens, rats and mice [19,20].

The life cycle of *Blastocystis* sp. remains elusive, although different morphological forms have been described, including vacuolar, granular, amoeboid and cysts. Recently, Tan [1] suggested a life cycle with the cyst as the infectious stage. After ingestion of cysts, the parasite may undergo excystation in the gastrointestinal tract and may develop into a vacuolar form that divides by binary fission. The following stage could be either the amoeboid form or the granular form. Then, encystation may occur during passage along the colon before cyst excretion in the feces. Therefore, *Blastocystis* sp. lives in oxygen-poor environments and is characterized by the presence of some double-membrane surrounded-organelles showing elongate, branched, and hooked cristae [21] called mitochondria-like organelles (MLOs) [22]. These cellular compartments contain a circular DNA molecule and have metabolic properties of both aerobic and anaerobic mitochondria [23,24].

Blastocystis sp. has been reported as a parasite causing gastro- and extra-intestinal diseases with additional persistent rashes, but a clear link of subtypes to the symptomatology is not well established [11]. Other studies have shown that the parasite can be associated with irritable bowel syndrome [20,25] or inflammatory bowel disease [26]. Thus, the pathogenic role of *Blastocystis* sp. as the primary cause of enteric symptoms is dubious. Therefore, it is important to search for other molecular markers for an epidemiologically integrated study [17]. Here we report the complete genome sequence of a subtype 7 isolate from a Singaporean patient [GenBank: CABX01000000]. Its comparison with the two other available stramenopile genome sequences (that is, *Phytophthora sojae*, a plant pathogenic oomycete, and *Thalassiosira pseudonana*, a free diatom) allows us to highlight some genome-specific features of *Blastocystis* to understand how this parasite evolved within

environmental constraints, but also provides a better knowledge of its metabolic and physiological capacities, such as the functioning and the role of MLOs and the arsenal produced to interact or to counter immune defense systems of its host.

Results and discussion

General features of the *Blastocystis* genome

The genome of a *Blastocystis* subtype 7 was resolved by pulsed-field gel electrophoresis, and 15 chromosomal bands have been characterized. The final assembled sequence is distributed in 54 scaffolds and the deduced genome is 18.8 Mb in size (16.5-fold sequence coverage), which is much smaller than plant parasite stramenopiles (*Phytophthora infestans*, 240 Mb; *P. sojae*, 95 Mb; *Phytophthora ramorum*, 65 Mb) and also smaller than free stramenopiles (*Phaeodactylum tricornutum*, 27.4 Mb; *T. pseudonana*, 34.5 Mb). The reference annotation of the *Blastocystis* subtype 7 genome contains 6,020 genes, covering about 42% of the genome (Table 1). The average number of exons per gene is 4.6 for multiexonic genes and 929 genes are monoexonic. Compaction in this parasite genome is reflected by the short length of the intergenic regions (1,801 bp), the relatively low repeat coverage (25%) and, more strikingly, by the very short size of introns, with a sharp length distribution of around 32 nucleotides (Figure S1 in Additional file 1). A total of 38 rDNA units organized in transcriptional units, including a small subunit rRNA gene, a 5.8S rRNA gene, and a large subunit rRNA gene in a 5'-3' orientation, have been detected in the genome. The sizes of the small subunit, the large subunit and the 5.8S rRNA gene are 1.8 kb, 2.45 kb and 0.44 kb, respectively. Some units are tandemly duplicated, up to four copies on scaffold 18, and some may also be localized in subtelomeric regions, as revealed by a co-mapping of telomeric sequences and rDNA subunits at scaffold 6 and 9 extremities. These two scaffolds could correspond to entire chromosomes. Due to the sequencing method, some units are incomplete (either truncated or lacking genes). The alignment of 20 complete small subunit rRNA genes shows polymorphism between copies, which is also the case for 29 large subunit rRNA gene copies.

The number of genes in *Blastocystis* (6,020) is reduced in comparison with other stramenopiles (*P. infestans*, 17,797; *P. sojae*, 19,027; *P. ramorum*, 15,743; *P. tricornutum*, 10,402; *T. pseudonana*, 11,776). Surprisingly, a large portion of genes were probably duplicated since 404 clusters of paralogous protein-coding genes were identified, containing 1,141 genes, that is, 19% of *Blastocystis* genes (see Material and methods). Excluding the large multigenic families (up to 32 genes with a histone-fold domain and 20 genes with a 4Fe-4S ferredoxin domain), most of the duplicated genes are present in only two copies (Figure S2 in

Table 1 General features of *Blastocystis* sp. subtype 7

	Number	Mean length	Median length	Total length (Mb)	Percentage of genome (18.8 Mb)
Genes	6,020	1,299	1,397	7.82	42%
Exons	24,580	280	150	6.88	37%
Introns	18,560	50.5	31	0.94	5%
Intergenic	-	1,801	4,092	10.9	58%
Repeats	2,730	1,747	2,862	4.8	25%

Additional file 1). As described in other organisms [27,28], the duplicated genes are more conserved than single copy genes in *Blastocystis* sp. Indeed, they have more orthologs (defined as best reciprocal hit (BRH); see Materials and methods) and display higher similarities with their orthologs (Figure S3 in Additional file 1). They also tend to display higher expression levels than single copy genes (Figure S4 in Additional file 1).

We investigated whether these gene duplications could have arisen from a whole genome duplication (WGD) or smaller scale segmental duplications. WGD, the duplication of the entire genome by polyploidization, has been shown to have played a key role in the evolutionary history of several animal and plant lineages [27,29-31]. Segmental duplications occur continually by several mechanisms that can duplicate parts of genes, entire genes, or several adjacent genes. These mechanisms include unequal crossing over, or gene conversion, and tandem duplication [32-34]. We were able to identify 320 blocks of duplicated genes, that is, paralogous segments of several adjacent genes (see Materials and methods), some of which are very large (up to 100 kb), suggesting a WGD. These blocks cover about 39% of the genome (7.3 out of 18.8 Mb) representing 38% (5.15 out of 13.65 Mb) of the unrepeat fraction of the genome. As shown in Figure 1, each scaffold is a mosaic of blocks of homology with several other scaffolds: scaffolds cannot be grouped by pairs as would be expected from a recent WGD. Additionally, some segments are present in more than two copies in the genome (they appear in black in Figure 1), suggesting that segmental duplications are likely to have played a role in the current duplication pattern. However, the duplicated blocks are not often on the same scaffold, nor in tandem, which rules out the tandem duplication model. The comparison of paralogous copies shows surprisingly high nucleic acid identity rates: on average, 99% in coding regions, 98.4% in untranslated regions, and 97.8% in introns and intergenic regions. Interestingly, those values are homogeneous among all paralogous blocks, suggesting that all blocks were duplicated at the same time.

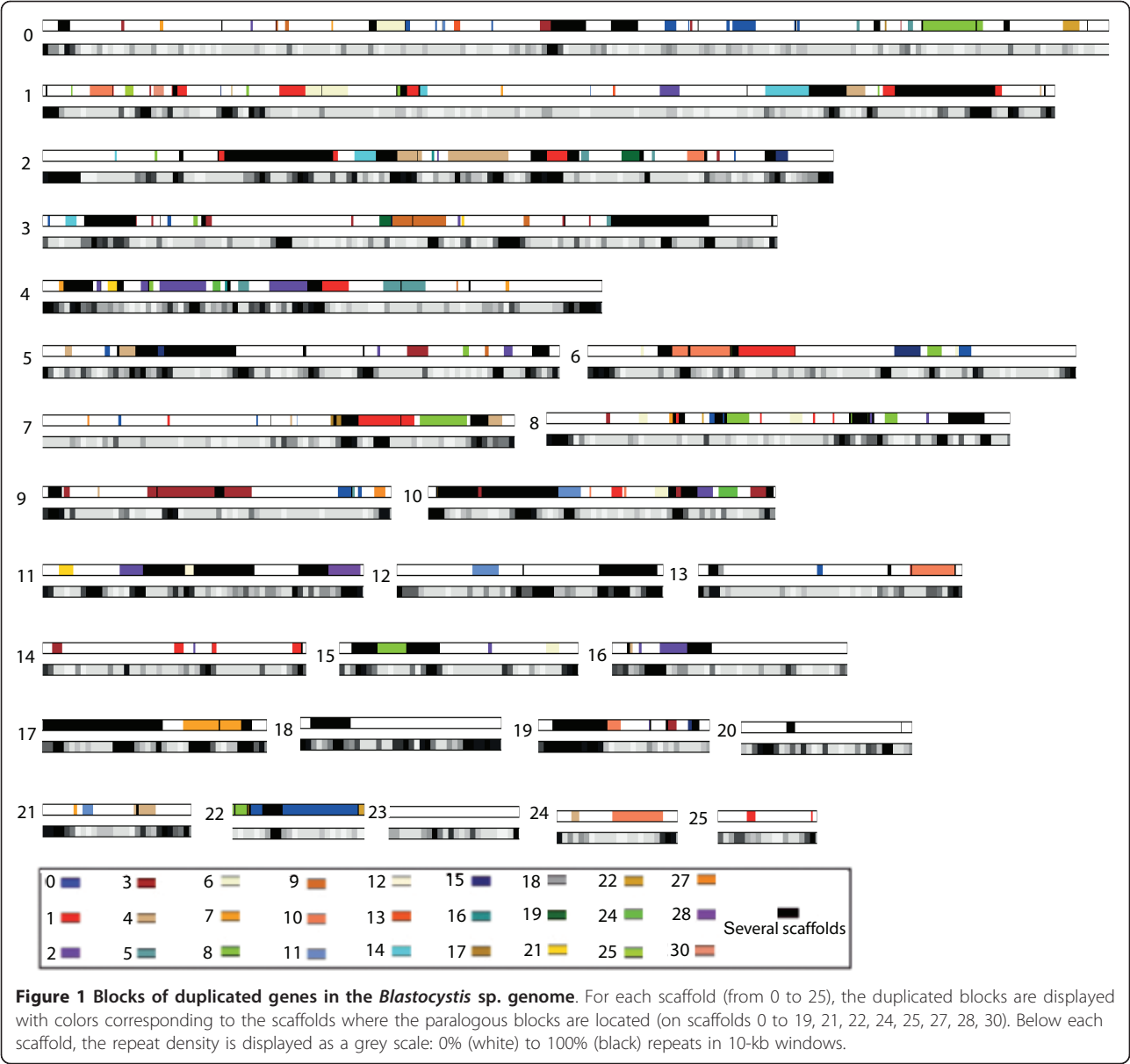
Two hypotheses could explain the origin of these duplicated blocks. First, the duplicates may have arisen from a whole genome duplication that took place recently (since

the copies are still very similar) and was followed by rapid genome rearrangements and losses of gene copies. The high homology between gene copies could also result from a high rate of homogenization through gene conversion driven by the high frequency of rearrangements. The frequent rearrangements in the *Blastocystis* lineage are probably also the reason why no extensive synteny could be detected between *Blastocystis* sp. and other stramenopiles. Second, the duplicates could also have occurred through segmental duplications (favored by the high rate of rearrangements), although the relatively uniform divergence between copies is more symptomatic of a single event and would imply a burst of segmental duplications during a short period or a very high rate of homogenization by recombination. The intriguing pattern of gene duplications, likely caused by the high rate of rearrangements in the *Blastocystis* genome, makes it impossible to determine which scenario is the most likely. It could be interesting to sequence other subtypes to determine whether the high rate of recombination (loss of synteny) and the pattern of duplications observed in subtype 7 is a common feature within this lineage.

Endosymbiotic and horizontal gene transfers in *Blastocystis* sp

Phylogenetic analyses revealed two genes of possible cyanobacterial origin in the genome of *Blastocystis*, those encoding phosphoglycerate kinase [GenBank: CBK20833] and 6-phosphogluconate dehydrogenase [GenBank:CBK22626] (Figure S5 in Additional file 1). It is important to notice that 6-phosphogluconate dehydrogenase-encoding genes have been identified in non-photosynthetic protists such as Heterolobosea (not shown). This was interpreted as secondary horizontal gene transfer (HGT) from photosynthetic eukaryotes to Heterolobosea [35,36].

The presence of plastids in various photosynthetic stramenopile lineages (for example, diatoms, chrysophytes, raphidophytes) was interpreted as a secondary endosymbiosis that occurred between a red algae and the ancestor of these groups. By contrast, the evolutionary meaning of the lack of plastids in some heterotrophic stramenopile lineages (for example, oomycetes, bicospesids) is still under discussion: does it indicate secondary losses of the



plastid acquired by the ancestor of all stramenopiles? Or does it reflect the fact that the secondary endosymbiosis at the origin of stramenopile plastids did not occur in their common ancestor but after the divergence of heterotrophic lineages [37]? The presence of genes of cyanobacterial origin in *Blastocystis* supports the first hypothesis even if we can not rule out possible recent acquisitions of genes of chloroplastic origin from photosynthetic eukaryotes as in the case of Heterolobosea.

HGT is important in evolution as an adaptive mechanism of microbial eukaryotes to environmental conditions [38,39] and is known to play an important role in stramenopiles. For instance, iron is a limiting nutrient in surface waters for diatoms. Therefore, the likely acquisition of

ferritin by HGT from bacteria has permitted some species to acquire this nutrient from the environment [40]. This is also the case for the diatom *Phaeodactylum*, in which nitrogen metabolism, cell wall silification, DNA replication, genome repair and recombination processes have been shaped by HGT [40,41]. HGT seems also to play an important role in oomycetes since it may be involved in osmotrophy. Genes involved in absorbing products of degradation of complex nutrients were predicted to be candidates for fungi-to-oomycete HGT [42]. By analyzing the set of predicted genes in *Blastocystis* sp. that are homologous to bacterial or archaeal genes, we identified 133 candidates for HGT (Table S3 in Additional file 2). In most cases, our phylogenetic analyses

confirm the bacterial origin of these genes even if they were not sufficiently resolved to allow the precise identification of the donor, suggesting that these HGT events were ancient and/or that the corresponding genes are rapidly evolving in the genome of *Blastocystis* sp. Interestingly, in a few cases, even when the transferred gene is of bacterial origin, the *Blastocystis* sp. copy is closely related to homologues found in pathogenic and/or anaerobic eukaryotes, suggesting that HGT between eukaryotes has played a key role in these organisms too (Figure S6 in Additional file 1).

Some of the genes that originated from HGT possess functions that lead to a better understanding of how this lineage emerged. Three are homologous to the bacterial major facilitator transporter (MFS_1), the first two being nearly identical, and therefore resulting from a recent gene duplication event. MFS proteins form a large and diverse group of secondary transporters, which facilitate the transport across membranes of a variety of substrates, including ions, sugar phosphates, drugs, neurotransmitters, nucleosides, amino acids and peptides [43]. Two *Blastocystis* MFS genes have closely related homologues in some pathogenic eukaryotes like the Alveolata *Perkinsus marinus* or fungi such as *Gibberella zeae* and *Verticillium albo atrum*, suggesting an acquisition from bacteria followed by HGT between these eukaryotes (Figure S6f in Additional file 1). However, the phylogeny resolution is too low to precisely identify the bacterial donor of these genes. The presence of MSF proteins in *Blastocystis* sp. may confer the ability to absorb nutrients from the environment to this parasite, particularly in the intestinal lumen or when attacking host tissues. We have also found different HGT genes harboring alcohol deshydrogenase, short-chain dehydrogenase and oxidoreductase domains (Table S3 in Additional file 2) that may be involved in specific fermentations that remain to be characterized. Some of them are closely related to homologues found in anaerobic eukaryotes like *Trichomonas vaginalis* and *Entamoeba histolytica* (Figure S6b in Additional file 1) or in the bacteria *Legionella pneumophila* or *Parachlamydia acanthamoebae*, which infect or are associated with amoeba [44,45]. These enzymes may increase the range of *Blastocystis* sp. metabolic abilities to produce energy in anaerobic environments, as has been observed in *Giardia lamblia* and *E. histolytica* [46,47].

Several genes acquired by HGT may participate in the adhesion of the parasite to the host tissues. Indeed, 26 genes (Table S3 in Additional file 2) encode proteins containing the IPR008009 domain, which is often associated with immunoglobulin domains, a conserved core region of an approximately 90-residue repeat found in several hemagglutinins and other cell surface proteins. Among these 26 *Blastocystis* sp. proteins, some also contain the IPR015919 domain, which characterizes

cadherins, a family of adhesion molecules that mediate Ca^{2+} -dependent cell-cell adhesion. Homologous genes are also found in some beta-Proteobacteria or Acidobacteria, but the sequences are very divergent and our phylogenetic analysis did not, therefore, allow firm identification of the bacterial donor. Some hydrolase-encoding genes could also result from the transfer from bacteria to *Blastocystis* sp. One of them possesses an esterase-lipase (IPR013094) domain (Table S3 in Additional file 2) and may participate in the degradation of host tissue during infection. The closest homologues of this gene are found in the fungus *Botryotinia fuckeliana*, in Firmicutes and Actinobacteria (Figure S6d in Additional file 1).

Overall, these HGT genes may have allowed flexibility in genome expression, enabling the successful adaptation of *Blastocystis* sp. to digestive environments through genes encoding proteins that could be involved in osmotrophy (MFS), energy metabolism (dehydrogenases) and adhesion.

Circular genome, predicted proteome and metabolic pathways of the MLOs

Although it lives in anaerobic or microaerophilic conditions, *Blastocystis* sp. harbors MLOs that present both mitochondrial and hydrogenosomal features [24]. We recently reported that *Blastocystis* sp. MLOs contain a circular genome, including genes encoding 10 of the 20 complex I subunits, but they lack all genes encoding cytochromes, cytochrome oxidases and ATP synthase subunits [24], unlike mitochondrial DNA from other sequenced stramenopiles, such as *Phytophthora* sp. [48]. The MLO genome of the *Blastocystis* subtype 7 is a circular molecule 29,270 bp in size. Two other MLO genomes were then sequenced from isolates belonging to other subtypes [49]: a subtype 1, represented by *Blastocystis* Nand II, with a 27,719 bp genome; and a subtype 4, represented by *Blastocystis* DMP/02-328, with a 28,382 bp genome. In addition to sequence conservation, these three genomes have many similarities. Their A+T content is around 80%, their gene density is higher than 95% and all three encompass 45 genes: 27 ORFs, 16 tRNAs and 2 rRNA genes. The ORFs consist of NADH subunits, ribosomal proteins and proteins with no similarity in the databases. The synteny between the three MLO genomes is highly conserved: gene order is strictly the same among the three genomes [24,49].

Through the analysis of a *Blastocystis* EST database, Stechmann et al. [23] have identified 110 potential proteins associated with mitochondrial pathways, such as the oxidative phosphorylation chain, tricarboxylic acid (TCA) cycle, Fe/S cluster assembly, and amino acid and fatty acid metabolisms. Nonetheless, approximately half of these proteins have an incomplete amino terminus

due to EST data, making it difficult to confirm mitochondrial import by algorithms. To clarify the metabolic characteristics of these puzzling organelles, we used data from the whole genome sequence in order to establish the *in silico* proteome of *Blastocystis* MLOs. For this purpose, a computational approach based on two different prediction algorithms (MitoProt and MitoPred) for mitochondrial-import proteins was chosen (see Materials and methods for more details). This approach predicted 365 MLO proteins (Table S6 in Additional file 3) whereas Stechmann *et al.* [23] predicted only 110 proteins. Among these 365 proteins, 299 were predicted to have an amino-terminal extension involved in mitochondrial import, suggesting that an alternative system might exist for the 66 remaining proteins. Of the 299 proteins, 41 remain as 'hypothetical protein' with unknown function and 31 have no homologues in public databases, which raises the question of the existence of undiscovered metabolic processes within these intriguing organelles (Table S6 in Additional file 3). The other proteins are involved in classical mitochondrial core functions, such as oxidative phosphorylation, amino acid metabolism, fatty acid oxidation, iron-sulfur cluster assembly, and mitochondrial import system. Several proteins involved in the translocase of the outer mitochondrial membrane (TOM complex), the translocase of the inner membrane (TIM complex), and the presequence translocase-associated motor (PAM complex), which perform protein transport into the matrix, were identified. Interestingly, the two essential subunits of the mitochondrial processing peptidase heterodimer (MPP α/β), essential for the cleavage of the targeting peptide, were also found [50].

Our analyses revealed that MLOs probably have three ways to make acetyl-CoA from pyruvate, supported by the presence of the pyruvate dehydrogenase complex, pyruvate:ferredoxin oxidoreductase and pyruvate:NADP⁺ oxidoreductase (an amino-terminal pyruvate:ferredoxin oxidoreductase domain fused to a carboxy-terminal NADPH-cytochrome P450 reductase domain) (Figure 2). *Euglena gracilis* mitochondria include this feature, which provides adaptability to various oxygen levels [51], and this might be to a lesser extent the case for *Blastocystis* sp. We have also identified the 20 subunits of the *Blastocystis* sp. MLO complex I (ten are encoded by the MLO genome and ten by nuclear genes). The four nuclear-encoded subunits of the mitochondrial respiratory chain complex II were detected and this complex could function in two ways (via succinate dehydrogenase or fumarate reductase) [52]. We did not identify any genes encoding complexes III and IV subunits or ATP synthase. However, we have found components of the TCA cycle, which was shown to be involved with complex II (fumarate reductase) in fumarate respiration

in parasitic helminths [52]. Interestingly, we identified a gene encoding a terminal oxidase, called alternative oxidase (AOX), which could be the terminal electron acceptor of complexes I and II (Figure 2), allowing adaptation to oxygen stress and maintaining the NADH/NAD balance, as has been suggested for *Cryptosporidium parvum* [53,54]. These data raise questions about the electron acceptor when complex II has succinate dehydrogenase or fumarate reductase activity, the quinone used in this process and the role of the proton gradient.

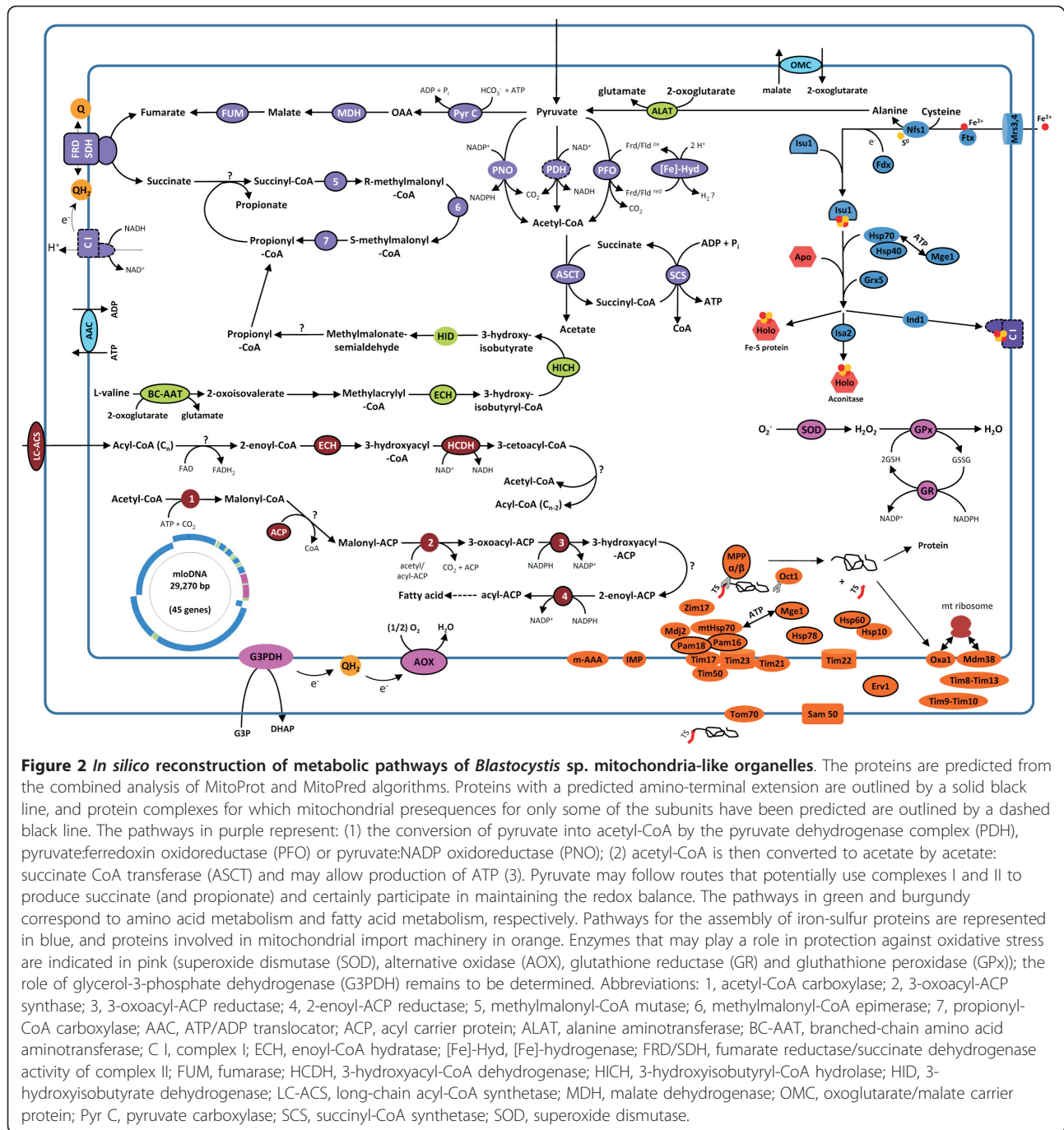
We also revealed proteins that can be grouped into essential mitochondrial pathways, like the Fe/S cluster assembly. More precisely, we have identified 11 enzymes (6 of which have predicted mitochondrial import signals), composing the iron-sulfur cluster system responsible for the assembly of mitochondrial Fe/S proteins [55], such as the cysteine desulfurase Nfs1, the scaffold protein Isu1, frataxin, and the P-loop NTPase Ind1, which is required for the assembly of complex I (Figure 2). We also highlighted some proteins involved in mitochondrial fatty acid synthesis type II [56], beta oxidation of fatty acids and amino acid metabolism (Table S6 in Additional file 3).

Taken together, our data confirm the mitochondrial nature of the *Blastocystis* sp. MLO. The oxygen-poor environment may have driven the selection of these unique organelles, which seemingly represent an intermediate situation between anaerobic mitochondria and hydrogenosomes, arguing for multiple situations arising during organelle evolution. It remains now to describe the metabolism occurring in these unusual organelles more precisely.

Secretome and virulence factors

The persistence of *Blastocystis* sp. in the host may be due, to some extent, to its ability to override the response of the immune system and to adhere and survive within the intestinal tissue. Manipulation of the host might be facilitated by molecules released at the interface between the host and the parasite [57]. Accordingly, the study of the predicted secretome of *Blastocystis* sp. is of particular interest. With SIGNALP 3.0, 307 proteins were predicted to be secretory, of which 46 had no sequence similarity in the public nr databases. By sequence homology, 170 proteins that could play a role in host-parasite relationships were selected and submitted to PSORTII for extracellular location. Finally, 75 putative secreted proteins have been classified by putative functions, some of which may have a direct connection with pathogenicity (proteases, hexose digestion enzymes, lectins, glycosyltransferases and protease inhibitors; Table S4 in Additional file 2).

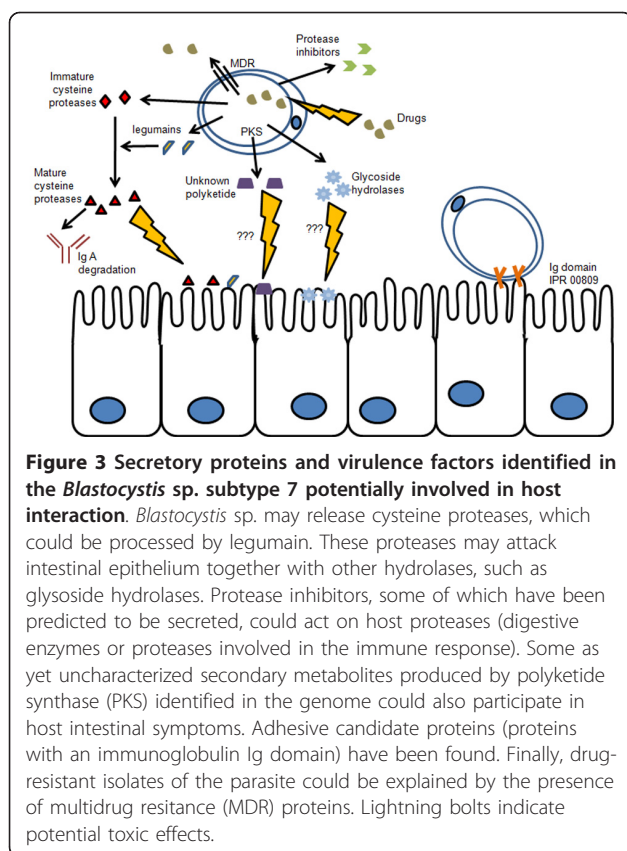
Blastocystis can secrete members of the immunophilin family, characterized by peptidyl-propyl *cis-trans*



isomerase activity and disulfide isomerases (Figure 3; Table S4 in Additional file 2). These proteins have key roles in protein folding, but it has also been established that they can have moonlighting functions. In bacteria, they have evolved adhesive properties for the host [58] but they can also modulate host leukocyte function and induce cellular apoptosis [59]. A cyclophilin-like protein from the protozoan parasite *Toxoplasma gondii* is directly involved in host-parasite crosstalk, as it can

modulate protective Th1 responses through its binding to the chemokine receptor CCR5 [60]. It is unclear what role these proteins play in *Blastocystis* sp., but this illustrates a range of functions for cell stress proteins in host-pathogen interactions.

Sugar-binding proteins have an important role through a conserved carbohydrate-recognition domain that could interact with host cell receptors. Such proteins have been characterized in other parasites [61] and



it is interesting to note that some sugar-binding proteins are able to inhibit Th1- and Th2-mediated inflammation [62,63]. Moreover, some specific sugar-binding proteins are also able to suppress regulatory T cells [64]. The binding of these proteins is dependent on their specific sugar motifs, which can be added to N- or O-linked glycans by glycosyltransferases. One carbohydrate-binding protein and eight glycosyltransferases (Table S4 in Additional file 2) have been predicted to be secreted. All these enzymes could allow cross-linking of *Blastocystis* sp. sugar-binding proteins to host cell receptors.

The parasite likely uses hydrolases to attack host tissues. Fucosidase, hexosaminidase and polygalacturonase have been identified in the predicted secretome and may participate in this process by degrading host glycoproteins (Figure 3; Table S4 in Additional file 2). Proteases have been proposed to be involved in diverse processes, such as host cell invasion, excystation, metabolism, cytoadherence or other virulence functions. A correlation between a high level of protease activity and the virulence of the intestinal parasite *E. histolytica* was proven by McKerrow *et al.* [65]. Indeed, cysteine proteases degrade extracellular matrix proteins, cleave immunoglobulin A and G, and are thought to be responsible for the cytopathic effect of different pathogens against *in vitro* cultured cells [66]. Interestingly, *Blastocystis* sp. proteolytic

enzymes are also able to degrade human secretory immunoglobulin A [67]. All the major classes of proteolytic enzymes were identified in the genome data, including serine, aspartic, and cysteine proteases and metalloproteases. Among the 66 proteases identified, 18 are predicted to be secreted by the parasite (Table S4 in Additional file 2). Within the protease family, cysteine protease-encoding genes are the most represented in *Blastocystis* sp. genome and 96% of the proteins encoded by these genes are predicted to be secreted. Among the cysteine proteases we have found five legumains and eight cathepsins; three cathepsins B contain the IPR015643 domain, which is only present in *Blastocystis* sp. compared to the other stramenopiles. The IPR015643 domain corresponds to the peptidase C1 cathepsin B domain and has a cysteine type peptidase activity, which was also found in pathogenic protozoa (*Leishmania* sp. and *Trypanosoma* sp.) [66]. Cysteine proteases are usually secreted in their inactive form and must be matured, having a prosegment that prevents hydrolysis during protease trafficking and storage. This maturation might result from the activity of the same protease or another, such as asparaginyl endopeptidase (also called legumain) [68]. This endopeptidase cleaves peptide bonds carboxy-terminal to asparagine residues, and may be involved in processing and activating both cathepsins L and B. Legumains have been predicted in the secretome of *Blastocystis* sp. (Table S4 in Additional file 2) and could be involved in protease processing (Figure 3). As an alternative role, secreted *Blastocystis* sp. legumains could also participate with other effectors in the alteration of the host intestine [69]. Indeed, it has been shown that legumain can degrade fibronectin, an extracellular matrix glycoprotein [70].

Genes coding for protease inhibitors are also present in the *Blastocystis* sp. genome, and some are predicted to be secreted. Release of protease inhibitors may weaken the host response as described in nematodes [71]. *Blastocystis* sp. encodes three protease inhibitors: cystatin, type1-proteinase inhibitor and endopeptidase inhibitor-like protein (Table S4 in Additional file 2). Type1-proteinase inhibitor is similar to chymotrypsin inhibitor, which is known to inactivate intestinal digestive enzymes (trypsin and chymotrypsin) as in *Ascaris suum* [72], thus protecting the parasite against non-specific digestive defenses. Cystatin, also called stefin, was described in *Fasciola gigantica* [73] and shown to inhibit mammalian cathepsin B, cathepsin L and other cysteine proteases, including parasite ones. In *Blastocystis* sp., secreted cystatin could participate in the regulation of parasitic cysteine protease activities. Cystatin can also potentially inhibit host proteases involved in MHC II antigen processing and presentation, including the key enzyme asparaginyl endopeptidase [74] and cathepsin S, the mammalian legumain [73].

Interestingly, a putative type I polyketide synthase (PKS) gene was also found in the *Blastocystis* sp. genome, potentially originating from HGT. PKS and non-ribosomal peptide synthetase (NRPS) synthesize metabolites like simple fatty acids, but also a myriad of chemical structures that possess important pharmacological activities and environmental impact, such as toxins, antibiotics or antimicrobials. Type I PKS was formerly known only from bacteria and fungi, but recently homologous genes were also discovered in some protists [75]. According to the Database for NRPS and PKS [76], the *Blastocystis* sp. PKS gene possesses the three essential domains, and three other domains: dehydratase, ketoacyl reductase, and enoyl reductase domains. The presence of these additional domains would permit this organism to synthesize both reduced polyketides and fatty acids. Domain comparison with other type I PKSs suggests that *Blastocystis* sp. PKS is similar to type I PKS from the ascomycete *Cochliobolus heterostrophus*, a maize pathogen that produces T toxin [77], a polyketide molecule that disturbs mitochondria by binding a protein of the inner mitochondrial membrane. Searching polyketide-related metabolites in the secretome of *Blastocystis* sp. would be of interest in order to identify molecules that could have effects on the host (Figure 3).

Antioxidant system and multi drug resistance

Like other anaerobic organisms, *Blastocystis* sp. has to eliminate reactive oxygen species such as superoxide anions ($O_2^{\cdot-}$), hydrogen peroxide (H_2O_2) and hydroxyl radicals (HO^{\cdot}) resulting from metabolism. In addition, this microorganism has to cope with the oxidative burst imposed by host immune cell effectors (release of $O_2^{\cdot-}$ subsequently processed to give additional reactive oxygen species). For these reasons, to protect against oxidative injury, *Blastocystis* species have developed an efficient battery of antioxidant enzymes (Table S5 in Additional file 2). The first lines of defense against oxygen damage are superoxide dismutases (SODs), a family of metalloproteins catalyzing the dismutation of $O_2^{\cdot-}$ to form H_2O_2 and oxygen. Genome annotation revealed the presence of two genes encoding SODs (SOD1 and SOD2) that exhibit sequence characteristics of dimeric iron-containing SODs [78] and likely protect the cytosol and MLOs, respectively, against $O_2^{\cdot-}$. Catalase and ascorbate peroxidase are subsequently able to remove H_2O_2 generated by SODs as well as by NADPH-dependent oxidase. However, genes encoding catalase and ascorbate peroxidase have not been identified in *Blastocystis* sp. nor in many unicellular parasites, including trypanosomatids and *Plasmodium falciparum*. Additional enzymes, glutathione peroxidase (Gpx) and thioredoxin-dependent peroxidase (commonly known as peroxyredoxin (Prx)) are able to reduce H_2O_2 to water

as well as other substrates, such as hydroperoxides and peroxynitrite. In most eukaryotes, both enzymes obtain their reducing equivalents from two redox systems, the glutathione (GSH) and the thioredoxin (Trx) systems, respectively. Like *P. falciparum* [79], *Blastocystis* sp. cells possess a complete GSH synthesis pathway: the genes encoding γ -glutamylcysteine synthetase, glutathione synthetase (eu-GS group) and a functional GSH/Gpx (nonselenium Gpx belonging to the PHGpx group)/glutathione reductase system have been identified and both Gpx and glutathione reductase are probably located in the MLO. This nearly ubiquitous redox cycle is replaced by the trypanothione system in trypanosomatids [80]. *Blastocystis* sp. also contains genes encoding the proteins of the Trx/thioredoxin reductase (TrxR)/Prx system. Indeed, two genes encode small proteins homologous to Trx: one cytosolic and another most likely located in the MLO (Table S5 in Additional file 2). Trx is itself reduced by TrxR and three genes encoding cytosolic TrxR have been identified in *Blastocystis* sp. These proteins clearly belong to the high molecular weight (designated H-TrxR) group of enzymes and are similar to metazoan enzymes, including those of *Homo sapiens* and *Drosophila melanogaster*, and to those of the apicomplexan protozoa *Plasmodium*, *Toxoplasma*, and *Cryptosporidium* [81]. Interestingly, in contrast to apicomplexan H-TrxRs, two of the H-TrxR enzymes of *Blastocystis* are predicted to possess a redox active center in the carboxy-terminal domain composed of a selenocysteine (a rare amino acid encoded by the opal codon TGA, which is not recognized as a stop codon) at the penultimate position and its neighboring cysteine residue as in metazoan enzymes (selenoprotein type H-TrxR). This strongly suggests the presence of the Se-Cys insertion machinery (SECYS elements) in *Blastocystis* sp. Genes encoding another type of TrxR with low molecular weight (designated L-TrxR) have been identified in parasitic protozoa such as *Trichomonas*, *Entamoeba*, and *Giardia* but not in the genome of *Blastocystis* sp. These data reinforce the assumption of the exclusive occurrence of either L-TrxR or H-TrxR in genomes and of some disadvantages of possessing both types of TrxR [81]. In *Blastocystis* sp., at least 11 highly similar gene copies encoding predicted cytosolic Prxs have been found that clearly belong to the typical 2-Cys class of Prx. Whether sequence polymorphism of these enzymes is potentially correlated with diversified expression or even function remains to be explored. Another gene encoding a typical 2-Cys Prx, likely located in the MLO, has been identified in this parasite. Interestingly, like the homologous sequence of another stramenopile, *P. infestans*, this latter protein is fused to Trx with a WCGKC motif. As described above, *Blastocystis* sp. possesses a whole array of antioxidant enzymes protecting both the cytosol and

MLO. As shown in Table S5 in Additional file 2, these enzymes have distinct phylogenetic origins and most of them probably originate from prokaryote HGT. These antioxidant proteins attract attention in unicellular parasites as they have important functions in host-parasite interactions and constitute new drug targets for the design of inhibitors. Indeed, genetic approaches have undoubtedly shown that some anti-oxidant enzymes are essential for the survival of different parasitic species [82-86].

Some genes coding for multi-drug resistance pump proteins have also been discovered in the *Blastocystis* sp. genome. There are two classes of multi-drug resistance genes: the first class corresponds to proteins that are energized by ATP hydrolysis; the second class includes proteins that mediate the drug efflux reaction with a proton or sodium ion gradient. Among the first class, 24 ABC transporter genes were found. In eukaryotes the main physiological function of ABC transporters is the export of endogenous metabolites and cytotoxic compounds [87] and eight families of ABC transporters (ABC A to H) have been identified. The *Blastocystis* sp. ABC transporters are included in four of these eight families (five in family A, six in family B, six in family C, three in family F, and four not in any class). The A family is involved in lipid trafficking, and the F family in DNA repair and gene regulation. The other two families are more interesting [87], since in protozoan parasites (*Leishmania* spp., *Trypanosoma* spp., *Plasmodium* spp.) transporters belonging to the B and C families confer resistance to drugs. Metronidazole-resistant strains of *Blastocystis* sp. could have arisen through the action of these multi-drug resistance proteins (Figure 3).

Conclusions

We have provided the first genome sequence of a *Blastocystis* sp. subtype, which could serve in comparative genomics studies with other subtypes to provide clues to clarify how these protozoans develop pathogenicity in some humans. Analysis of this genome has revealed original traits of this lineage compared to other stramenopiles (free living and plant pathogens). Aerobic respiration has been lost, *Blastocystis* sp. instead having the MLO, an anaerobic organelle, which should advance our understanding of organelle evolution as the *Blastocystis* sp. MLO seems to be unique among organelles (Figure 2) but remains to be biochemically characterized. Some genes may have been gained through HGT, which may participate in essential functions for an intestinal parasite (adhesion, energy production). These genes probably have facilitated adaptation to intestinal environments. The *Blastocystis* sp. secretome has been predicted and this has permitted the identification of candidate proteins that could degrade host tissues in

order to provide nutrients. Putative secretory proteins that can interfere with non-specific and specific host defense systems have also been found, enabling *Blastocystis* sp. to survive within this hostile environment (Figure 3). These putative secretory proteins are of particular interest as they may interact directly with host tissue and could help in understanding the host-parasite interactions and could also be used as markers to distinguish between non-pathogenic and pathogenic isolates. If their functions are essential, they could also be used to develop future vaccine formulations. The antioxidant proteins offer interesting therapeutic targets as they might be important for the parasite in fighting oxidative bursts. In summary, the deciphering of the *Blastocystis* sp. genome will contribute to the study of interactions between this parasite and its host at a post-genomic scale and pave the way for deciphering the host-parasite interactome. Finally, the '*Blastocystis* sp. story' is reminiscent of the amoeba pathogenicity story where two morphologically indistinguishable species have different pathogenic potential [88], and this genome will help in the development of typing tools for the characterization of pathogenic isolates.

Materials and methods

Genome sequencing

The *Blastocystis* sp. genome was sequenced using a whole genome shotgun strategy. All data were generated by paired-end sequencing of cloned inserts using Sanger technology on ABI3730xl sequencers. Table S1 in Additional file 2 gives the number of reads obtained per library. All reads were assembled with Arachne [89]. We obtained 157 contigs that were linked into 54 supercontigs. The contig N50 was 297 kb, and the supercontig N50 was 901 kb (Table S2 in Additional file 2).

Genome annotation

Construction of the training set

A set of 300 gene models from a preliminary annotation run was selected randomly, among those that were validated by *Blastocystis* sp. cDNAs (that is, with every intron confirmed by at least one cDNA and no exon overlapping a cDNA intron) to create a clean *Blastocystis* sp. training set. This training set was used to train gene prediction algorithms and optimize their parameters.

Repeat masking

Most of the genome comparisons were performed with repeat masked sequences. For this purpose, we searched and masked sequentially several kinds of repeats: known repeats and transposons available in Repbase with the Repeat Masker program [90], tandem repeats with the TRF program [91], *ab initio* repeat detection with RepeatScout [92], rDNA by BLATing [93] 189 rDNAs

sequences (downloaded from GenBank), and telomeric repeats by searching 'CCCTAA' patterns in the scaffolds with the BLAST2 algorithm.

GeneWise

The UniProt [94] database was used to detect conserved genes between *Blastocystis* sp. and other species. As GeneWise [95] is time greedy, the UniProt database was first aligned with the *Blastocystis* sp. genome assembly using BLAT [93]. Subsequently, we extracted the genomic regions where no protein hit had been found by BLAT and realigned Uniprot proteins with more permissive parameters. Each significant match was then refined using GeneWise in order to identify exon/intron boundaries.

GeneID and SNAP

GeneID [96] and SNAP [97] *ab initio* gene prediction software were trained on 300 genes from the training set.

Blastocystis sp. cDNAs

Full-length-enriched cDNA libraries were constructed from *Blastocystis* sp. vacuolar forms using a SV total RNA isolation system (Promega France, Charbonnières, France) for RNA extraction. RNA quality and quantity were estimated using the Agilent bioanalyser with the RNA 6000 Nano LabChip[®] Kit. The clones were sequenced on the 5' end, producing 34,470 useful reads. We were able to align 33,685 cDNA sequences to the *Blastocystis* sp. genome assembly with the following pipeline: after masking of polyA tails, the sequences were aligned with BLAT on the assembly and all matches with scores within 99% of the best score were extended by 5 kb on each end, and realigned with the cDNA clones using the EST2genome software [98].

Stramenopile ESTs

A collection of 410,069 public mRNAs from the stramenopile clade (276,208 downloaded from the National Center for Biotechnology Information plus 43,932 and 80,929 ESTs downloaded from the Joint Genome Institute for diatoms and *Ectocarpus*, respectively) were first aligned with the *Blastocystis* sp. genome assembly using BLAT [93]. To refine BLAT alignment, we used EST2genome [98]. Each significant match was chosen for an alignment with EST2genome. BLAT alignments were made using default parameters between translated genomic and translated ESTs.

Integration of resources using GAZE

All the resources described here were used to automatically build *Blastocystis* sp. gene models using GAZE [99]. Individual predictions from each of the programs (that is, GeneID, SNAP, GeneWise, EST2genome) were broken down into segments (coding, intron, intergenic) and signals (start codon, stop codon, splice acceptor, splice donor, transcript start, transcript stop).

Exons predicted by *ab initio* software (that is, GeneWise and EST2genome) were used as coding segments. Introns predicted by GeneWise and EST2genome were used as intron segments. Intergenic segments were created from the span of each mRNA using a negative score (coercing GAZE not to split genes). Predicted repeats were used as intron and intergenic segments to avoid prediction of genes coding proteins in such regions.

The whole genome was scanned to find signals (splice sites and start and stop codons). Additionally, transcript stop signals were extracted from the ends of mRNAs (polyA tail positions).

Each segment extracted from software output that predicts exon boundaries (like GeneWise, Exonerate or *ab initio* predictors) was used by GAZE only if GAZE chose the same boundaries. Each segment or signal from a given program was given a value reflecting our confidence in the data, and these values were used as scores for the arcs of the GAZE automaton. All signals were given a fixed score, but segment scores were context sensitive: coding segment scores were linked to the percentage identity of the alignment; intronic segment scores were linked to the percentage identity of the flanking exons. A weight was assigned to each resource to further reflect its reliability and accuracy in predicting gene models. This weight acts as a multiplier for the score of each information source, before processing by GAZE. When applied to the entire assembled sequence, GAZE predicted 4,798 gene models. Since the resource of expressed sequences in stramenopiles is limited, and some gene-free 'holes' appeared in gene-dense regions, we suspected that some genes had been missed by the annotation pipeline because of a lack of support.

Additional gene models

With the assumption that not all genes in *Blastocystis* sp. have EST support, we developed the following strategy to recuperate additional gene models. *Ab initio* (SNAP and GeneID) predictions that did not overlap GAZE gene models were selected and aligned to UniProt sequences. Predictions that had significant hits (coverage $\geq 90\%$; e-value $\leq 10^{-5}$) were tagged as potential coding genes and randomly chosen genes were successfully verified by RT-PCR using the Access RT-PCR system (Promega France, Charbonnières, France). The final proteome composed of 6,020 gene models was obtained by adding 1,222 supplementary models to the 4,798 genes from the first GAZE output.

Identification of orthologous genes

We identified orthologous genes with three species: *Cyandioschyzon merolae* [100], *P. sojae* [49] and *T. pseudonana* [101]. Each pair of predicted genes was aligned with the Smith-Waterman algorithm, and alignments with a score

higher than 300 (BLOSUM62, gapo = 10, gape = 1) were retained. Orthologs were defined as BRHs, that is, two genes, A from genome GA and B from genome GB, were considered orthologs if B is the best match for gene A in GB and A is the best match for B in GA.

Identification of paralogous genes and duplicated blocks

An all-against-all comparison of *Blastocystis* sp. proteins was performed using the Smith-Waterman algorithm implemented in the Biofacet package [102]. BRHs were identified as follows: two genes, A and B, are the BRH if B is the best match for gene A and A for gene B. The distribution of percentage identities among the pairs of BRHs is displayed in Figure S7 in Additional file 1. The distribution is widespread except for the abundant class of genes sharing $\geq 90\%$ of identity, which represents 48% of all pairs of paralogs. We investigated this apparently recent gene duplication by selecting all pairs of genes sharing $\geq 90\%$ identity over $\geq 50\%$ of the length of the shortest protein (not only BRHs), which gave 1,917 gene pairs corresponding to 1,141 genes scattered in 404 gene families (19% of *Blastocystis* sp. genes). The number of counterparts per gene is displayed in Figure S2 in Additional file 1. Additionally, blocks of paralogous genes, or so-called duplicated blocks, were identified by clustering the 1,917 gene pairs. The clustering was performed by single linkage clustering using the Euclidian distance between genes, and independently of gene orientation. Those distances were calculated with the gene index on each scaffold rather than the genomic position, including only the genes with paralogs. The minimal distance between two paralogous genes was set to 5 and the minimal number of genes in a cluster was set to 4 (two pairs of paralogous genes; Figure S8 in Additional file 1).

Identification of candidate horizontal gene transfers

Blastocystis sp. proteins were blasted [103] (blastx) against the protein nr database with the parameters '-f 100 -X 100 -e 0.00001 -E 2 -W 5', and the best hits were retained using the following criteria: for BLAST scores greater than 200, all hits with a score greater than 90% of the best score were retained; and for BLAST scores lower or equal to 200, all hits with a score greater than 80% of the best score were retained. Then, the proteins with all their best hits in bacteria or archaea were retained as candidates that had potentially arisen from HGT. Other criteria for the blastx comparison were tested (such as W = 3) but we observed no significant difference in the results after the subsequent filters. Candidates with some of their best hits in stramenopiles in addition to bacteria were also retained since some HGTs may be shared between stramenopiles, and genes for which orthologs were identified in non-stramenopile

species were discarded. The evolutionary origin of the candidate genes was then investigated using phylogenetic approaches (Figure S6 in Additional file 1). For each gene, homologues were retrieved from the protein nr database using Blastp (default parameters, except for the max-target-sequences threshold, which was fixed at 500). The sequences were aligned using Muscle 3.6 [104] (default parameters). The resulting alignments were visually inspected and manually refined using the MUST software [105]. Ambiguously aligned regions were removed prior to phylogenetic analysis.

Maximum likelihood phylogenetic tree reconstructions were carried out on the remaining positions using PhyML [106] with the Le and Gascuel (LG) model [106] with a gamma correction (four discrete classes, an estimated alpha parameter) to take into account evolutionary rate variation among sites. Tree robustness was estimated by a non-parametric bootstrap approach using PhyML and the same parameters with 100 replicates of the original dataset. Bayesian phylogenetic trees were also reconstructed using MrBayes version 3.1.2 [107]. We used a mixed model of amino acid substitution and a gamma distribution (four discrete categories plus a proportion of invariant sites) to take into account site rate variation. MrBayes was run with four chains for 1 million generations and trees were sampled every 100 generations. To construct the consensus tree, the first 1,500 trees were discarded as 'burn-in'. The candidates with clear eukaryotic origin were then discarded. This process provided 133 candidate genes (Table S3 in Additional file 2). These candidates contain a high proportion of monoexonic genes (39%) compared to the average number of monoexonic genes in *Blastocystis* sp. (approximately 15%).

Protein domain analysis

InterProScan [108] was run against all *C. merolae*, *P. sojae*, *T. pseudonana* and *Blastocystis* sp. proteins. Matches that fulfilled the following criteria were retained: match tagged as 'true positive' by InterProScan (status = T); match with an e-value $\leq 10^{-1}$. A total of 2,305 InterPro domains (with IPR number) were found in *Blastocystis* sp., which corresponds to 4,096 proteins.

Functional annotation

Enzyme annotation

Enzyme detection in predicted *Blastocystis* sp. proteins was performed with PRIAM [109], using the PRIAM July 2006 Enzyme release. A total of 428 different EC numbers, corresponding to enzyme domains, are associated with 1,140 *Blastocystis* sp. proteins. Therefore, about 19% of *Blastocystis* sp. proteins contain at least one enzymatic domain.

Association of metabolic pathways with enzymes and *Blastocystis* sp

Potential metabolic pathways were deduced from EC numbers using the KEGG pathway database [110]. Links between EC numbers and metabolic pathways were obtained from the KEGG website. Using this file and the PRIAM results, 906 (of the 1,140) *Blastocystis* sp. proteins were assigned to 201 pathways.

Identification of putative proteins imported within the MLOs

The whole proteome was scanned by two algorithms aimed at predicting proteins imported to mitochondria; MitoProt [111], which predicts mitochondrial-targeting sequences, and MitoPred [112], which predicts nuclear-encoded mitochondrial proteins based on Pfam domains (animal/yeast database). After manual processing and using a script, only protein sequences with a score above 0.5 and 85% for MitoProt and MitoPred, respectively, were selected. This output file was then used in a KEGG Automatic Annotation Server (KAAS) with the bi-directional best hit method [113] in order to automatically generate KEGG pathways. Because protein domain annotations did not always provide sufficient information (PRIAM July 2006 Enzyme release), a BLAST comparison against the non-redundant database was conducted.

Secretome prediction using SignalP 3.0 and pSORTII

Prediction of secreted proteins is based on the analysis of amino-terminal secretory signal sequences (SignalP 3.0) followed by the selection of proteins predicted as extracellular by pSORTII. Each of the proteins was individually submitted to SignalP 3.0 for analysis with the following parameters: organism set to eukaryotes, output short format and protein sequence truncation after the first 50 amino acids. Results of SignalP 3.0 were exported to a temporary file, and identification of signal peptides was accomplished by parsing the results of the hidden Markov model analysis conducted by SignalP 3.0. Proteins with secretory signals were retained and analyzed on the basis of possible function in host-parasite interactions. These last ones were also analyzed using PSORT II [114], and those having a best hit as 'extracellular' were selected. The SignalP threshold value for secretory signal peptide prediction was set at 0.5 as determined for previous analyses [115] and the best hit was chosen for the PSORTII analysis. The predicted secretory proteins were then annotated as functional protein families.

Additional material

Additional file 1: Genome organization of *Blastocystis* sp. (introns, numbers of counterparts per gene, genome structure, and so on) and phylogenetic trees illustrating horizontal gene transfer events from prokaryotic donors to *Blastocystis* sp. and candidate genes for endosymbiotic gene transfers of chloroplastic origin.

Additional file 2: Sequencing overview and assembly metric data, and the identification of horizontal gene transfer, secretory protein and antioxidant protein candidates.

Additional file 3: Proteins putatively imported in the mitochondria-like organelle.

Abbreviations

bp: base pair; BRH: best reciprocal hit; EST: expressed sequence tag; Gpx: glutathione peroxidase; GSH: glutathione; HGT: horizontal gene transfer; KEGG: Kyoto Encyclopedia of Genes and Genomes; MFS: major facilitator transporter; MLO: mitochondria-like organelle; NRPS: non-ribosomal peptide synthase; ORF: open reading frame; PKS: polyketide synthase; Prx: peroxiredoxin; SOD: superoxide dismutase; TCA: tricarboxylic acid; Trx: thioredoxin; TrxR: thioredoxin reductase; WGD: whole genome duplication.

Acknowledgements

We would like to thank François Enault (Université Blaise Pascal) for SignalP 3.0 analysis, David G Biron (Université Blaise Pascal) and Susan Cure (Genoscope, Evry) for manuscript reading, comments and English corrections.

Author details

¹Genoscope (CEA) and CNRS UMR 8030, Université d'Evry, 2 rue Gaston Crémieux, 91057 Evry, France. ²Clermont Université, Université Blaise Pascal, Laboratoire Microorganismes: Génome et Environnement, BP 10448, F-63000 Clermont-Ferrand, France. ³CNRS, UMR 6023, LMGE, F-63177 Aubière, France. ⁴Center for Infection and Immunity of Lille, Institut Pasteur de Lille, F-59019 Lille Cedex, France. ⁵Inserm U1019, F-59000 Lille Cedex, France. ⁶CNRS UMR 8402, F-59021 Lille Cedex, France. ⁷University Lille-Nord de France, F-59000 Lille Cedex, France. ⁸Laboratoire de chimie bactérienne (CNRS UPR9043), Institut de Microbiologie de la Méditerranée, 31 chemin Joseph Aiguier, 13402 Marseille, France. ⁹Université de Provence, Aix-Marseille I, 3 place Victor Hugo, 13331 Marseille, France. ¹⁰Laboratory of Molecular and Cellular Parasitology, Department of Microbiology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, 5 Science Drive 2, 117597 Singapore.

Authors' contributions

FDen, PW, CPV and HEA conceived and designed the experiments. MR, IW, JP, GCN, BS, BN, CDS, AC and HEA performed the experiments. MR, IW, MD, CT, BN, EV, CBA, FDen, VA, FA, JMA, OJ, KSWT, FDel, PW and HEA analyzed the data. FDen, MR, IW, EV, CBA, FDel, CPV and HEA wrote the paper. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 25 October 2010 Revised: 4 January 2011

Accepted: 25 March 2011 Published: 25 March 2011

References

1. Tan KS: New insights on classification, identification, and clinical relevance of *Blastocystis* spp. *Clin Microbiol Rev* 2008, **21**:639-665.
2. Alexeieff A: Sur la nature des formations dites "kystes de *Trichomonas intestinalis*". *CR Soc Biol* 1911, **71**:296-298.
3. Silberman JD, Sogin ML, Leipe DD, Clark CG: Human parasite finds taxonomic home. *Nature* 1996, **380**:398.
4. Arisue N, Hashimoto T, Yoshikawa H, Nakamura Y, Nakamura G, Nakamura F, Yano TA, Hasegawa M: Phylogenetic position of *Blastocystis hominis* and of stramenopiles inferred from multiple molecular sequence data. *J Eukaryot Microbiol* 2002, **49**:42-53.
5. Hoevers JD, Snowden KF: Analysis of the ITS region and partial ssu and lsu rRNA genes of *Blastocystis* and *Proteromonas lacertae*. *Parasitology* 2005, **131**:187-196.
6. Patterson DJ: The diversity of eukaryotes. *Am Nat* 1999, **154**:S96-S124.
7. Brumpt E: *Blastocystis hominis* n. sp. et formes voisines. *Bull Soc Pathol Exot* 1912, **5**:725-730.

8. Windsor JJ, Macfarlane L, Hughes-Thapa G, Jones SK, Whiteside TM: **Incidence of *Blastocystis hominis* in faecal samples submitted for routine microbiological analysis.** *Br J Biomed Sci* 2002, **59**:154-157.
9. Stark D, van Hal S, Marriott D, Ellis J, Harkness J: **Irritable bowel syndrome: a review on the role of intestinal protozoa and the importance of their detection and diagnosis.** *Int J Parasitol* 2007, **37**:11-20.
10. Rivera WL: **Phylogenetic analysis of *Blastocystis* isolates from animal and human hosts in the Philippines.** *Vet Parasitol* 2008, **156**:178-182.
11. Souppart L, Moussa H, Cian A, Sancier G, Poirier P, El Alaoui H, Delbac F, Boorom K, Delhaes L, Dei-Cas E, Viscogliosi E: **Subtype analysis of *Blastocystis* isolates from symptomatic patients in Egypt.** *Parasitol Res* 2010, **106**:505-511.
12. Marciano MG, Takizawa H, Falavigna DLM, Gomes ML: **Enteroparasitosis and their ethnographic relationship to food handlers in a tourist and economic center in Paraná, southern Brazil.** *Rev Inst Med trop S Paulo* 2009, **51**:31-35.
13. Souppart L, Sancier G, Cian A, Wawrzyniak I, Delbac F, Capron M, Dei-Cas E, Boorom K, Delhaes L, Viscogliosi E: **Molecular epidemiology of human *Blastocystis* isolates in France.** *Parasitol Res* 2009, **105**:413-421.
14. Cirioni O, Giacometti A, Drenaggi D, Ancarani F, Scalise G: **Prevalence and clinical relevance of *Blastocystis hominis* in diverse patient cohorts.** *Eur J Epidemiol* 1999, **15**:389-393.
15. Lucia JF, Aguilar C, Betran A: ***Blastocystis hominis* colitis in a haemophilic patient as a cause of lower gastrointestinal bleeding.** *Haemophilia* 2007, **13**:224-225.
16. Stenzel DJ, Boreham PF: ***Blastocystis hominis* revisited.** *Clin Microbiol Rev* 1996, **9**:563-584.
17. Stensvold CR, Suresh GK, Tan KS, Thompson RC, Traub RJ, Viscogliosi E, Yoshikawa H, Clark CG: **Terminology for *Blastocystis* subtypes - a consensus.** *Trends Parasitol* 2007, **23**:93-96.
18. Wong KH, Ng GC, Lin RT, Yoshikawa H, Taylor MB, Tan KS: **Predominance of subtype 3 among *Blastocystis* isolates from a major hospital in Singapore.** *Parasitol Res* 2008, **102**:663-670.
19. Iguchi A, Ebisu A, Nagata S, Saitou Y, Yoshikawa H, Iwatani S, Kimata I: **Infectivity of different genotypes of human *Blastocystis hominis* isolates in chickens and rats.** *Parasitol Int* 2007, **56**:107-112.
20. Boorom KF, Smith H, Nimri L, Viscogliosi E, Spanakos G, Parkar U, Li LH, Zhou XN, Ok UZ, Leelayoova S, Jones MS: **Oh my aching gut: irritable bowel syndrome, *Blastocystis*, and asymptomatic infection.** *Parasit Vectors* 2008, **1**:40.
21. Zierdt CH: ***Blastocystis hominis*-past and future.** *Clin Microbiol Rev* 1991, **4**:61-79.
22. Nasirudeen AMA, Eu-Hian Y, Singh M, Tan KSW: **Metronidazole induces programmed cell death in the protozoan parasite *Blastocystis hominis*.** *Microbiology* 2004, **150**:33-43.
23. Stechmann A, Hamblin K, Perez-Brocal V, Gaston D, Richmond GS, van der Giezen M, Clark CG, Roger AJ: **Organelles in *Blastocystis* that blur the distinction between mitochondria and hydrogenosomes.** *Curr Biol* 2008, **18**:580-585.
24. Wawrzyniak I, Roussel M, Diogon M, Couloux A, Texier C, Tan KS, Vivares CP, Delbac F, Wincker P, El Alaoui H: **Complete circular DNA in the mitochondria-like organelles of *Blastocystis hominis*.** *Int J Parasitol* 2008, **38**:1377-1382.
25. Windsor JJ: ***Blastocystis hominis* and *Dientamoeba fragilis*: neglected human protozoa.** *The Biomedical Scientist* 2007, **64**:524-527.
26. al-Tawil YS, Gilger MA, Gopalakrishna GS, Langston C, Bommer KE: **Invasive *Blastocystis hominis* infection in a child.** *Arch Pediatr Adolesc Med* 1994, **148**:882-885.
27. Jaillon O, Aury JM, Brunet F, Petit JL, Stange-Thomann N, Mauceli E, Bouneau L, Fischer C, Ozouf-Costaz C, Bernot A, Nicaud S, Jaffe D, Fisher S, Lutfalla G, Dossat C, Segurens B, Dasilva C, Salanoubat M, Levy M, Boudet N, Castellano S, Anthouard V, Jubin C, Castelli V, Katinka M, Vacherie B, Biémont C, Skalli Z, Cattolico L, Poulain J, et al: **Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype.** *Nature* 2004, **431**:946-957.
28. Davis JC, Petrov DA: **Preferential duplication of conserved proteins in eukaryotic genomes.** *PLoS Biol* 2004, **2**:E55.
29. Wolfe KH, Shields DC: **Molecular evidence for an ancient duplication of the entire yeast genome.** *Nature* 1997, **387**:708-713.
30. Semple C, Wolfe KH: **Gene duplication and gene conversion in the *Caenorhabditis elegans* genome.** *J Mol Evol* 1999, **48**:555-564.
31. Vision TJ, Brown DG, Tanksley SD: **The origins of genomic duplications in *Arabidopsis*.** *Science* 2000, **290**:2114-2117.
32. Sankoff D: **Gene and genome duplication.** *Curr Opin Genet Dev* 2001, **11**:681-684.
33. Bailey JA, Gu Z, Clark RA, Reinert K, Samonte RV, Schwartz S, Adams MD, Myers EW, Li PW, Eichler EE: **Recent segmental duplications in the human genome.** *Science* 2002, **297**:1003-1007.
34. Wolfe KH, Li WH: **Molecular evolution meets the genomics revolution.** *Nat Genet* 2003, **33**(Suppl):255-265.
35. Andersson JO, Roger AJ: **A cyanobacterial gene in nonphotosynthetic protists - an early chloroplast acquisition in eukaryotes?.** *Curr Biol* 2002, **12**:115-119.
36. Maruyama S, Misawa K, Iseki M, Watanabe M, Nozaki H: **Origins of a cyanobacterial 6-phosphogluconate dehydrogenase in plastid-lacking eukaryotes.** *BMC Evol Biol* 2008, **8**:151.
37. Archibald JM: **The puzzle of plastid evolution.** *Curr Biol* 2009, **19**:R81-88.
38. Keeling PJ, Palmer JD: **Horizontal gene transfer in eukaryotic evolution.** *Nat Rev Genet* 2008, **9**:605-618.
39. Andersson JO: **Horizontal gene transfer between microbial eukaryotes.** *Methods Mol Biol* 2009, **532**:473-487.
40. Keeling PJ: **Functional and ecological impacts of horizontal gene transfer in eukaryotes.** *Curr Opin Genet Dev* 2009, **19**:613-619.
41. Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, Kuo A, Maheswari U, Martens C, Maumus F, O'tillar RP, Rayko E, Salamov A, Vandepoele K, Beszteri B, Gruber A, Heijde M, Katinka M, Mock T, Valentin K, Verret F, Berges JA, Brownlee C, Cadoret JP, Chiovitti A, Choi CJ, Coesel S, De Martino A, Detter JC, Durkin C, Falciatore A, et al: **The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes.** *Nature* 2008, **456**:239-244.
42. Richards TA, Dacks JB, Jenkinson JM, Thornton CR, Talbot NJ: **Evolution of filamentous plant pathogens: gene exchange across eukaryotic kingdoms.** *Curr Biol* 2006, **16**:1857-1864.
43. Law CJ, Maloney PC, Wang DN: **Ins and outs of major facilitator superfamily antiporters.** *Annu Rev Microbiol* 2008, **62**:289-305.
44. Berger P, Papazian L, Drancourt M, La Scola B, Auffray JP, Raoult D: **Ameba-associated microorganisms and diagnosis of nosocomial pneumonia.** *Emerg Infect Dis* 2006, **12**:248-255.
45. Embley TM, van der Giezen M, Horner DS, Dyal PL, Foster P: **Mitochondria and hydrogenosomes are two forms of the same fundamental organelle.** *Philos Trans R Soc Lond B Biol Sci* 2003, **358**:191-201, discussion 201-192.
46. Field J, Rosenthal B, Samuelson J: **Early lateral transfer of genes encoding malic enzyme, acetyl-CoA synthetase and alcohol dehydrogenases from anaerobic prokaryotes to *Entamoeba histolytica*.** *Mol Microbiol* 2000, **38**:446-455.
47. Nixon JE, Wang A, Field J, Morrison HG, McArthur AG, Sogin ML, Loftus BJ, Samuelson J: **Evidence for lateral transfer of genes encoding ferredoxins, nitroreductases, NADH oxidase, and alcohol dehydrogenase 3 from anaerobic prokaryotes to *Giardia lamblia* and *Entamoeba histolytica*.** *Eukaryot Cell* 2002, **1**:181-190.
48. Martin FN, Bensasson D, Tyler BM, Boore JL: **Mitochondrial genome sequences and comparative genomics of *Phytophthora ramorum* and *P. sojae*.** *Curr Genet* 2007, **51**:285-296.
49. Perez-Brocal V, Clark CG: **Analysis of two genomes from the mitochondrion-like organelle of the intestinal parasite *Blastocystis*: complete sequences, gene content, and genome organization.** *Mol Biol Evol* 2008, **25**:2475-2482.
50. Chacinska A, Koehler CM, Milenkovic D, Lithgow T, Pfanner N: **Importing mitochondrial proteins: machineries and mechanisms.** *Cell* 2009, **138**:628-644.
51. Hoffmeister M, van der Klei A, Rotte C, van Grinsven KW, van Hellemond JJ, Henze K, Tielens AG, Martin W: ***Euglena gracilis* rhodoquinone:ubiquinone ratio and mitochondrial proteome differ under aerobic and anaerobic conditions.** *J Biol Chem* 2004, **279**:22422-22429.
52. Tielens AG, Rotte C, van Hellemond JJ, Martin W: **Mitochondria as we don't know them.** *Trends Biochem Sci* 2002, **27**:564-572.
53. Putignani L, Tait A, Smith HV, Horner D, Tovar J, Tetley L, Wastling JM: **Characterization of a mitochondrion-like organelle in *Cryptosporidium parvum*.** *Parasitology* 2004, **129**:1-18.
54. Henriquez FL, Richards TA, Roberts F, McLeod R, Roberts CW: **The unusual mitochondrial compartment of *Cryptosporidium parvum*.** *Trends Parasitol* 2005, **21**:68-74.

55. Lill R: **Function and biogenesis of iron-sulphur proteins.** *Nature* 2009, **460**:831-838.
56. Hiltunen JK, Schonauer MS, Autio KJ, Mittelmeier TM, Kastaniotis AJ, Dieckmann CL: **Mitochondrial fatty acid synthesis type II: more than just fatty acids.** *J Biol Chem* 2009, **284**:9011-9015.
57. Corrales RM, Sereno D, Mathieu-Daude F: **Deciphering the *Leishmania* exoproteome: what we know and what we can learn.** *FEMS Immunol Med Microbiol* 2010, **58**:27-38.
58. Bell A, Monaghan P, Page AP: **Peptidyl-prolyl cis-trans isomerases (immunophilins) and their roles in parasite biochemistry, host-parasite interaction and antiparasitic drug action.** *Int J Parasitol* 2006, **36**:261-276.
59. Henderson B: **Cell stress proteins as modulators of bacteria - host interactions.** *Novartis Found Symp* 2008, **291**:141-154, discussion 154-149, 221-144.
60. Golding H, Aliberti J, King LR, Manischewitz J, Andersen J, Valenzuela J, Landau NR, Sher A: **Inhibition of HIV-1 infection by a CCR5-binding cyclophilin from *Toxoplasma gondii*.** *Blood* 2003, **102**:3280-3286.
61. Klion AD, Donelson JE: **OvGalBP, a filarial antigen with homology to vertebrate galactoside-binding proteins.** *Mol Biochem Parasitol* 1994, **65**:305-315.
62. Toscano MA, Commodaro AG, Ilarregui JM, Bianco GA, Liberman A, Serra HM, Hirabayashi J, Rizzo LV, Rabinovich GA: **Galectin-1 suppresses autoimmune retinal disease by promoting concomitant Th2- and T regulatory-mediated anti-inflammatory responses.** *J Immunol* 2006, **176**:6323-6332.
63. Katoh S, Ishii N, Nobumoto A, Takeshita K, Dai SY, Shinonaga R, Niki T, Nishi N, Tominaga A, Yamauchi A, Hirashima M: **Galectin-9 inhibits CD44-hyaluronan interaction and suppresses a murine model of allergic asthma.** *Am J Respir Crit Care Med* 2007, **176**:27-35.
64. Kubach J, Lutter P, Bopp T, Stoll S, Becker C, Huter E, Richter C, Weingarten P, Warger T, Knop J, Müller S, Wjdenes J, Schild H, Schmitt E, Jonuleit H: **Human CD4+CD25+ regulatory T cells: proteome analysis identifies galectin-10 as a novel marker essential for their anergy and suppressive function.** *Blood* 2007, **110**:1550-1558.
65. McKerrow JH, Sun E, Rosenthal PJ, Bouvier J: **The proteases and pathogenicity of parasitic protozoa.** *Annu Rev Microbiol* 1993, **47**:821-853.
66. Klemm M, Goldberg DE: **Biological roles of proteases in parasitic protozoa.** *Annu Rev Biochem* 2002, **71**:275-305.
67. Puthia MK, Vaithilingam A, Lu J, Tan KS: **Degradation of human secretory immunoglobulin A by *Blastocystis*.** *Parasitol Res* 2005, **97**:386-389.
68. Dalton JP, Brindley PJ, Donnelly S, Robinson MW: **The enigmatic asparaginyl endopeptidase of helminth parasites.** *Trends Parasitol* 2009, **25**:59-61.
69. Wu B, Yin J, Texier C, Roussel M, Tan KS: ***Blastocystis* legumain is localized on the cell surface, and specific inhibition of its activity implicates a pro-survival role for the enzyme.** *J Biol Chem* 2010, **285**:1790-1798.
70. Morita Y, Araki H, Sugimoto T, Takeuchi K, Yamane T, Maeda T, Yamamoto Y, Nishi K, Asano M, Shirahama-Noda K, Nishimura M, Uzu T, Hara-Nishimura I, Koya D, Kashiwagi A, Ohkubo I: **Legumain/asparaginyl endopeptidase controls extracellular matrix remodeling through the degradation of fibronectin in mouse renal proximal tubular cells.** *FEBS Lett* 2007, **581**:1417-1424.
71. Hartmann S, Lucius R: **Modulation of host immune responses by nematode cystatins.** *Int J Parasitol* 2003, **33**:1291-1302.
72. Martzen MR, Peanasky RJ: ***Ascaris suum*: biosynthesis and iso-inhibitor profile of chymotrypsin/elastase iso-inhibitors.** *Exp Parasitol* 1985, **59**:313-320.
73. Tarasuk M, Vichasri Grams S, Vyanant V, Grams R: **Type I cystatin (stefin) is a major component of *Fasciola gigantica* excretion/secretion product.** *Mol Biochem Parasitol* 2009, **167**:60-71.
74. Manoury B, Gregory WF, Maizels RM, Watts C: **Bm-CPI-2, a cystatin homolog secreted by the filarial parasite *Brugia malayi*, inhibits class II MHC-restricted antigen processing.** *Curr Biol* 2001, **11**:447-451.
75. Smith S, Tsai SC: **The type I fatty acid and polyketide synthases: a tale of two megasynthases.** *Nat Prod Rep* 2007, **24**:1041-1072.
76. Database for NRPS and PKS. [http://www.nii.res.in/searchall.html].
77. Yang G, Rose MS, Turgeon BG, Yoder OC: **A polyketide synthase is required for fungal virulence and production of the polyketide T-toxin.** *Plant Cell* 1996, **8**:2139-2150.
78. Wintjens R, Noel C, May AC, Gerbod D, Dufernez F, Capron M, Viscogliosi E, Rooman M: **Specificity and phenetic relationships of iron- and manganese-containing superoxide dismutases on the basis of structure and sequence comparisons.** *J Biol Chem* 2004, **279**:9248-9254.
79. Muller S: **Redox and antioxidant systems of the malaria parasite *Plasmodium falciparum*.** *Mol Microbiol* 2004, **53**:1291-1305.
80. Muller S, Liebau E, Walter RD, Krauth-Siegel RL: **Thiol-based redox metabolism of protozoan parasites.** *Trends Parasitol* 2003, **19**:320-328.
81. Hirt RP, Muller S, Embley TM, Coombs GH: **The diversity and evolution of thioredoxin reductase: new perspectives.** *Trends Parasitol* 2002, **18**:302-308.
82. Krnajska Z, Gilberger TW, Walter RD, Cowman AF, Muller S: **Thioredoxin reductase is essential for the survival of *Plasmodium falciparum* erythrocytic stages.** *J Biol Chem* 2002, **277**:25970-25975.
83. Krieger S, Schwarz W, Ariyanayagam MR, Fairlamb AH, Krauth-Siegel RL, Clayton C: **Trypanosomes lacking trypanothione reductase are avirulent and show increased sensitivity to oxidative stress.** *Mol Microbiol* 2000, **35**:542-552.
84. Plewes KA, Barr SD, Gedamu L: **Iron superoxide dismutases targeted to the glycosomes of *Leishmania chagasi* are important for survival.** *Infect Immun* 2003, **71**:5910-5920.
85. Wilkinson SR, Horn D, Prathalingam SR, Kelly JM: **RNA interference identifies two hydroperoxide metabolizing enzymes that are essential to the bloodstream form of the african trypanosome.** *J Biol Chem* 2003, **278**:31640-31646.
86. Piacenza L, Alvarez MN, Peluffo G, Radi R: **Fighting the oxidative assault: the *Trypanosoma cruzi* journey to infection.** *Curr Opin Microbiol* 2009, **12**:415-421.
87. Sauvage V, Aubert D, Escotte-Binet S, Villena I: **The role of ATP-binding cassette (ABC) proteins in protozoan parasites.** *Mol Biochem Parasitol* 2009, **167**:81-94.
88. Pinilla AE, Lopez MC, Viasus DF: **[History of the *Entamoeba histolytica* protozoan].** *Rev Med Chil* 2008, **136**:118-124.
89. Jaffe DB, Butler J, Gnerre S, Mauceli E, Lindblad-Toh K, Mesirov JP, Zody MC, Lander ES: **Whole-genome sequence assembly for mammalian genomes: Arachne 2.** *Genome Res* 2003, **13**:91-96.
90. Chen N: **Using RepeatMasker to identify repetitive elements in genomic sequences.** *Curr Protoc Bioinformatics* 2004, **Chapter 4**, Unit 4 10.
91. Benson G: **Tandem repeats finder: a program to analyze DNA sequences.** *Nucleic Acids Res* 1999, **27**:573-580.
92. Price AL, Jones NC, Pevzner PA: **De novo identification of repeat families in large genomes.** *Bioinformatics* 2005, **21**(Suppl 1):i351-358.
93. Kent WJ: **BLAT - the BLAST-like alignment tool.** *Genome Res* 2002, **12**:656-664.
94. Bairoch A, Apweiler R, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh LS: **The Universal Protein Resource (UniProt).** *Nucleic Acids Res* 2005, **33**:D154-159.
95. Birney E, Durbin R: **Using GeneWise in the *Drosophila* annotation experiment.** *Genome Res* 2000, **10**:547-548.
96. Parra G, Blanco E, Guigo R: **GenelD in *Drosophila*.** *Genome Res* 2000, **10**:511-515.
97. Korf I: **Gene finding in novel genomes.** *BMC Bioinformatics* 2004, **5**:59.
98. Mott R: **EST_GENOME: a program to align spliced DNA sequences to unspliced genomic DNA.** *Comput Appl Biosci* 1997, **13**:477-478.
99. Howe KL, Chothia T, Durbin R: **GAZE: a generic framework for the integration of gene-prediction data by dynamic programming.** *Genome Res* 2002, **12**:1418-1427.
100. Nozaki H, Takano H, Misumi O, Terasawa K, Matsuzaki M, Maruyama S, Nishida K, Yagisawa F, Yoshida Y, Fujiwara T, Takio S, Tamura K, Chung SJ, Nakamura S, Kuroiwa H, Tanaka K, Sato N, Kuroiwa T: **A 100%-complete sequence reveals unusually simple genomic features in the hot-spring red alga *Cyanidioschyzon merolae*.** *BMC Biol* 2007, **5**:28.
101. Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou S, Allen AE, Apt KE, Bechner M, Brzezinski MA, Chaal BK, Chiovitti A, Davis AK, Demarest MS, Detter JC, Glavina T, Goodstein D, Hadi MZ, Hellsten U, Hildebrand M, Jenkins BD, Jurka J, Kapitonov VV, Kröger N, Lau WW, Lane TW, Larimer FW, Lippmeier JC, Lucas S, et al: **The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism.** *Science* 2004, **306**:79-86.
102. **GenomeQuest.** [http://www.gene-it.com].
103. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215**:403-410.

104. Edgar RC: **MUSCLE: a multiple sequence alignment method with reduced time and space complexity.** *BMC Bioinformatics* 2004, **5**:113.
105. Philippe H: **MUST, a computer package of Management Utilities for Sequences and Trees.** *Nucleic Acids Res* 1993, **21**:5264-5272.
106. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O: **New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0.** *Syst Biol* 2010, **59**:307-321.
107. Ronquist F, Huelsenbeck JP: **MrBayes 3: Bayesian phylogenetic inference under mixed models.** *Bioinformatics* 2003, **19**:1572-1574.
108. Zdobnov EM, Apweiler R: **InterProScan - an integration platform for the signature-recognition methods in InterPro.** *Bioinformatics* 2001, **17**:847-848.
109. Claudel-Renard C, Chevalet C, Faraut T, Kahn D: **Enzyme-specific profiles for genome annotation: PRIAM.** *Nucleic Acids Res* 2003, **31**:6633-6639.
110. Kanehisa M, Goto S, Kawashima S, Nakaya A: **The KEGG databases at GenomeNet.** *Nucleic Acids Res* 2002, **30**:42-46.
111. MITOPROT. [<http://ihg2.helmholtz-muenchen.de/ihg/mitoprot.html>].
112. MITOPRED. [<http://bioapps.rit.albany.edu/MITOPRED/>].
113. KEGG Automatic Annotation Server. [http://www.genome.jp/kaas-bin/kaas_main].
114. PSORT II Prediction. [<http://psort.ims.u-tokyo.ac.jp/form2.html>].
115. Nagaraj SH, Gasser RB, Ranganathan S: **Needles in the EST haystack: large-scale identification and analysis of excretory-secretory (ES) proteins in parasitic nematodes using expressed sequence tags (ESTs).** *PLoS Negl Trop Dis* 2008, **2**:e301.

doi:10.1186/gb-2011-12-3-r29

Cite this article as: Denoeud *et al.*: Genome sequence of the stramenopile *Blastocystis*, a human anaerobic parasite. *Genome Biology* 2011 **12**:R29.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



Rapid Communication

Complete circular DNA in the mitochondria-like organelles of *Blastocystis hominis*

Ivan Wawrzyniak^{a,1}, Michaël Roussel^{a,1}, Marie Diogon^a, Arnaud Couloux^b, Catherine Texier^a, Kevin S.W. Tan^c, Christian P. Vivarès^a, Frédéric Delbac^a, Patrick Wincker^b, Hicham El Alaoui^{a,*}

^a Laboratoire Microorganismes: Génome et Environnement, UMR CNRS 6023, Université Blaise Pascal, 63177 Aubière Cedex, France

^b Genoscope (CEA) and CNRS UMR 8030, Université d'Evry, 2 rue Gaston Crémieux CP5706, 91057 Evry, France

^c Laboratory of Molecular and Cellular Parasitology, Department of Microbiology, Yong Loo Lin School of Medicine, National University of Singapore, 5 Science Drive 2, 117597, Singapore

Received 23 April 2008; received in revised form 16 June 2008; accepted 17 June 2008

Abstract

Blastocystis hominis is an anaerobic parasite of the human intestinal tract belonging to the Stramenopile group. Using genome sequencing project data, we describe here the complete sequence of a 29,270-bp circular DNA molecule that presents mitochondrial features (such as oxidative phosphorylation complex I subunits) but lacks complexes III, IV and V. Transmission electron microscopy analyses reveal that this molecule, as well as mitochondrial (NADH:ubiquinone oxidoreductase subunit 7 (NAD7), β -succinyl-CoA synthetase (β -SCS)) and hydrogenosomal (pyruvate ferredoxin oxido-reductase (PFOR), iron-hydrogenase) proteins, are located within double-membrane surrounded-compartments known as mitochondria-like organelles (MLOs). As there is no evidence for hydrogen production by this organism, we suggest that MLOs are more likely anaerobic mitochondria.

© 2008 Australian Society for Parasitology Inc. Published by Elsevier Ltd. All rights reserved.

Keywords: *Blastocystis hominis*; MLOs; Circular DNA molecule; Iron-hydrogenase

Stramenopiles (Patterson, 1999) constitute a highly diverse group that includes free photosynthetic organisms such as diatoms, as well as non-photosynthetic taxa such as plant (*Phytophthora* spp.) and animal (*Blastocystis* spp.) parasites. They can live either in aerobic or in anaerobic conditions depending on their habitat (parasitic, freshwater, marine or terrestrial). *Blastocystis hominis* is a common anaerobic protozoan parasite found in the intestinal tract of humans and animals, that has recently been suggested to be involved in Irritable Bowel Syndrome (Stark et al., 2007). It harbours double-membrane surrounded-organelles of about 1 μ m, named “mitochondria-like organelles” (MLOs) (Nasirudeen and Tan, 2004), but their role and nature remain unclear. Most of

the microorganisms that live in different microaerophilic or anaerobic environments (e.g. protists, fungi and parasitic helminths), lack classical mitochondria but possess mitochondria-related organelles known as anaerobic mitochondria, hydrogenosomes or mitosomes. The two former are still able to produce energy, however mitosomes have lost this property but retain the iron–sulfur cluster assembly machinery (Hackstein et al., 2006). Although the origin of these organelles is controversial, most evidence supports evolution from mitochondria (Embley and Martin, 2006). A recent analysis of a *Blastocystis* expressed sequence tag (EST) database notably blurs the distinction between mitochondria and hydrogenosomes (Stechmann et al., 2008).

Our present study reports the first data from the whole genome sequencing project of *B. hominis* that is currently running at the Genoscope (Evry, France). The sequenced isolate belonged to the subtype 7 according to *Blastocystis* nomenclature standardization (Stensvold et al., 2007). It

* Corresponding author. Tel.: +33 4 73 40 74 31; fax: +33 4 73 40 76 70.
E-mail address: hicham.el_alaoui@univ-bpclermont.fr (H. El Alaoui).

¹ These authors made an equal contribution to the manuscript.

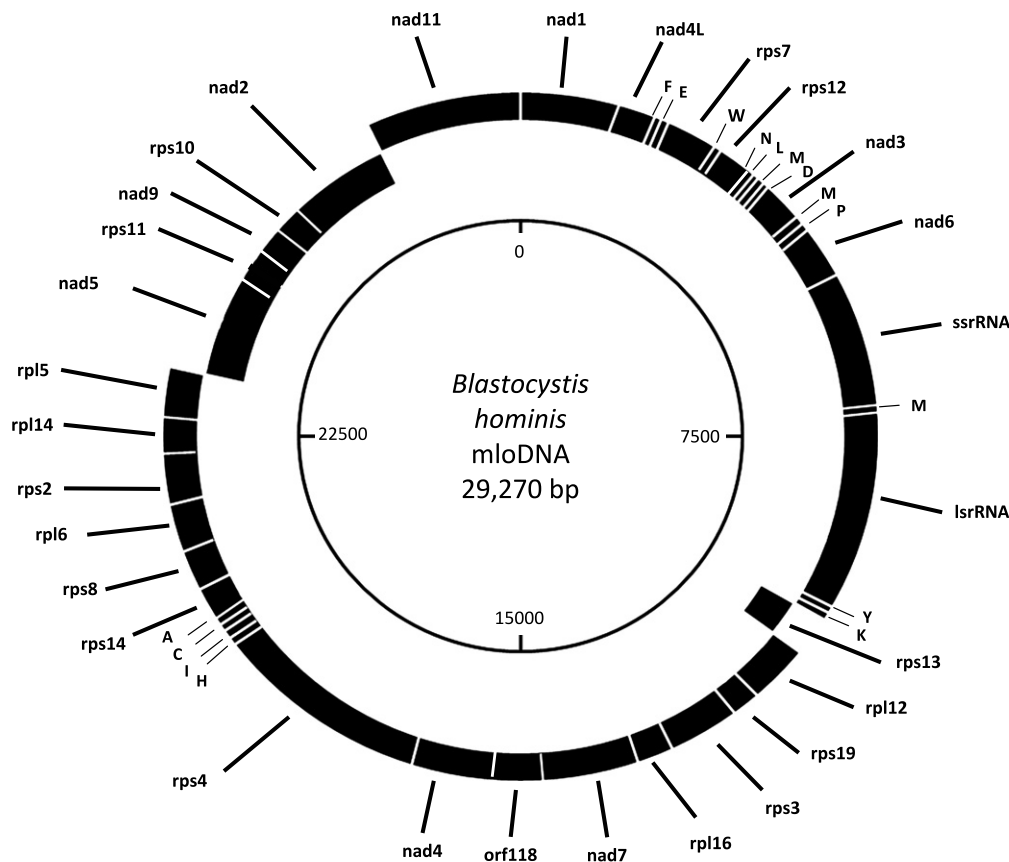


Fig. 1. *Blastocystis hominis* 29,270-bp circular DNA molecule (GenBank Accession No. CU914152) is of mitochondrial origin. Genes are represented by filled rectangles. Clockwise- and anti-clockwise-transcribed genes are located on the outermost and innermost circles, respectively. tRNA genes are indicated by the amino acid they carry. *Blastocystis hominis* circular DNA has a number of genes with significant similarity to mitochondrial genes. These correspond to 10 subunits of the NADH:ubiquinone oxidoreductase (nad), to five large (rpl) and 11 small (rps) ribosomal proteins, as well as ssrRNA and lsrRNAs. Finally, one open reading frame (Orf118) shares no sequence homology. When compared with *Phytophthora* spp. mitochondrial DNAs, which range from 37,957 for *Phytophthora infestans* (GenBank Accession No. U17009) to 42,977 bp for *Phytophthora sojae* (GenBank Accession No. DQ832717), the *B. hominis* circular DNA molecule has lost genes coding for components of oxidative phosphorylation complexes III (cob), IV (cox1, cox2, cox3) and V (atp1, atp6, atp8, atp9), and for the SecY independent transporter. Stechmann et al. (2008) have characterised a ~6-kbp sequence from *Blastocystis* sp. mitochondria-like organelles (MLOs), which corresponds to the position 2691–8293 on our sequence, thus from the tRNA asparagine gene to the last third of the lsrDNA one.

was obtained from a sample stool of a Singapore patient and maintained axenically and anaerobically as previously described (Ho et al., 1993). Total DNA was extracted from parasites that were included (1.10^7 cells/mL) in low melting point Agar block plugs (Seaplaque, Tebu-Bio) and treated by proteinase K (1 mg/mL, 48 h). Cloning was realised in pcDNA2.1, pCNS and pBBC plasmid vectors, and shotgun sequencing was performed using ABI3730xl analysers and BigDye V3.1 chemistry (Applied Biosystems). The reads were assembled with Arachne algorithm (Broad Institute). During the assembly step, two contigs clustered together forming a 29,270-bp circular DNA (Fig. 1). Artemis software (Sanger Institute, <http://www.sanger.ac.uk/Software/Artemis/>) and the BLAST network service (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) were, respectively, used to draw the DNA molecule open reading frames (ORFs) and to search for sequence similarities between those and the National Center for Biotechnology Information gene database (GenBank, <http://www.ncbi.nlm.nih.gov/>).

The *B. hominis* circular-assembled DNA presented mitochondrial characteristics (Barberà et al., 2007). Its overall A + T content was 79.96%, despite a lower A + T content observed in the lsrRNA gene (48.72%). Gene density was high with 94.78% of the entire sequence corresponding to coding DNA. Intergenic regions ranged from 1 to 397 nucleotides, with an average length of 42 nucleotides, and only four exceeding 100 nucleotides. Genome compactness was also seen with six gene overlaps. The *B. hominis* circular genome encompassed 45 genes: 16 tRNA and two rRNA genes, 26 ORFs with an assigned function, and one ORF with no detectable similarity in GenBank. Ten subunits of NADH:ubiquinone oxidoreductase (oxidative phosphorylation complex I) and most ribosomal proteins were easily identified. The conserved domain characterising some ribosomal proteins (RPS 3, 14, 19 and RPL 6, 14) was incomplete in the *B. hominis* counterparts. While the *B. hominis* RPS4 protein was longer than the other known RPS4 proteins and retained the family conserved domain,

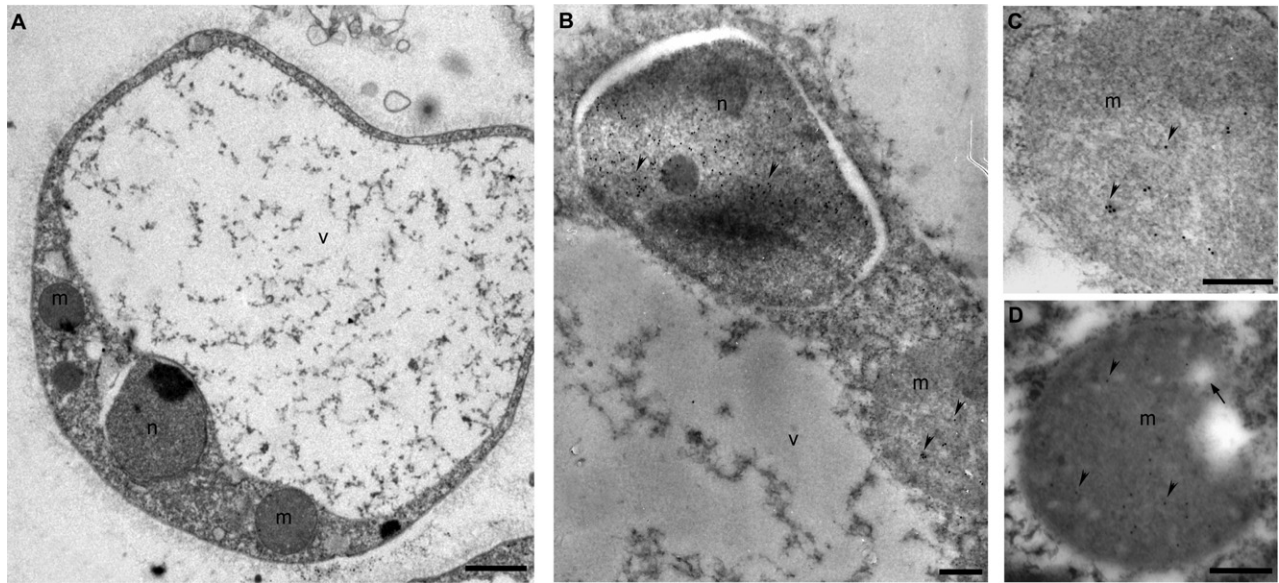


Fig. 2. The 29,270-bp circular DNA molecule is located within *B. hominis* mitochondria-like organelles (MLOs). (A) Electron micrograph of the in vitro cultivated vacuolar form of *B. hominis*. (B–C) Immunostaining of the vacuolar form using anti-DNA antibodies. Both the nucleus (n) and the MLOs (m) are labelled by the antibodies (arrowheads). C corresponds to a higher magnification of the MLO seen in (B). (D) Electron microscopic in situ hybridization (EM-ISH) with a biotinylated-probe designed from the *ssrRNA* sequence of the circular DNA molecule. Anti-biotin antibodies specifically label the MLO (arrowheads) confirming that the circular DNA is located within these organelles. Note the presence of cristae-like structures in the matrix of some MLOs (arrows). v, vacuole. Bars, 1 μ m for A and 400 nm for B–D. *Associated methods*: For transmission electron microscopy (TEM), in vitro cultivated *B. hominis* cells were embedded in Epikote resin (Agar Scientific, A) or in Unicryl resin (British BioCell International, B–D) following the procedure used by Mazet et al. (2008). In B and C, ultrathin sections were incubated for 3 h with an anti-DNA antibody (Santa Cruz Biotechnology, dilution 1:10) and for 1 h with a 1:100 dilution of goat anti-mouse IgG conjugated with 10 nm colloidal gold particles (Sigma). The *ssrRNA* probe (5'-CCTCTGTATTACCGCGGCTGCTGG-3') was synthesised by solid-phase phosphoramidite chemistry and labelled by tailing 3' ends with biotin (Eurogentec). The ultrathin section in D was incubated with this probe (1 μ M) for 3 h at 37 °C. Staining was done using goat anti-biotin antibodies (1:50, Sigma) and anti-goat secondary antibodies (1:50, Sigma) conjugated with 5 nm colloidal gold particles as described by Taupin et al. (2006). After staining, ultrathin sections were examined under a JEOL 1200 \times transmission electron microscope.

the RPS 2, 7 and RPL 5 proteins exhibited very low identities with orthologs. Nevertheless, the overall ribosomal gene organisation was conserved between *B. hominis* circular DNA and mtDNA from other sequenced Stramenopiles, such as *Phytophthora* spp. (Martin et al., 2007). The complete circular-assembled DNA of *B. hominis* covered the ~6000-bp fragment obtained by Stechmann et al. (2008) with a strictly conserved gene order. However, within this common region, the identity was around 87% because the DNA originated from two different isolates of *Blastocystis*.

It was previously suggested that *B. hominis* MLOs contained DNA (Nasirudeen and Tan, 2004). We confirmed this by immunoelectron microscopy with an anti-DNA antibody (Fig. 2B and C). The staining was observed not only in the nucleus but also in the MLOs matrix showing that these double-membrane surrounded-organelles contained DNA. The nucleus was more intensely labelled, suggesting DNA concentration discrepancies between both compartments. To demonstrate that the mlo DNA corresponded to the 29,270-bp molecule, electron microscopy in situ hybridization (EM-ISH) assays were carried out on parasites using a specific oligonucleotidic probe targeting the *ssrRNA*. As shown in Fig. 2D, this probe specifically labelled the MLOs, demonstrating their transcription machinery activity.

As the presence of a circular genome in double-membrane surrounded-organelles is a mitochondrial feature, *B. hominis* genome data were scanned for typical mitochondrial genes, i.e. genes coding for proteins involved in oxidative phosphorylation. We thus identified genes encoding 14 subunits of complex I, 10 on the circular mloDNA and four on the nuclear genome. These 14 subunits are considered to be the minimal set required for the bacterial complex I functionality, as exemplified by the *Escherichia coli* NDH-1 complex. Three functional modules can be distinguished in this complex (Vogel et al., 2007): (i) a dehydrogenase module responsible for the oxidation of NADH, (ii) an hydrogenase module guiding the released electrons, and (iii) a proton translocation or transporter module. Moreover, the four subunits of the succinate dehydrogenase complex (complex II) were also found in the *B. hominis* nuclear genome. In contrast, none of the genes coding the ubiquinol:cytochrome *c* oxidase complex (complex III), the cytochrome *c* oxidase complex (complex IV), or the ATP synthase (complex V) were identified. *Blastocystis hominis* genome data also contained sequences encoding an iron-hydrogenase and a pyruvate ferredoxin oxido-reductase (PFOR) both of these being enzymes typically recovered in hydrogenosomes. Finally, we also identified α and β subunits of succinyl-CoA synthetase (SCS), involved in energy production both in hydrogenosomes and mitochondria.

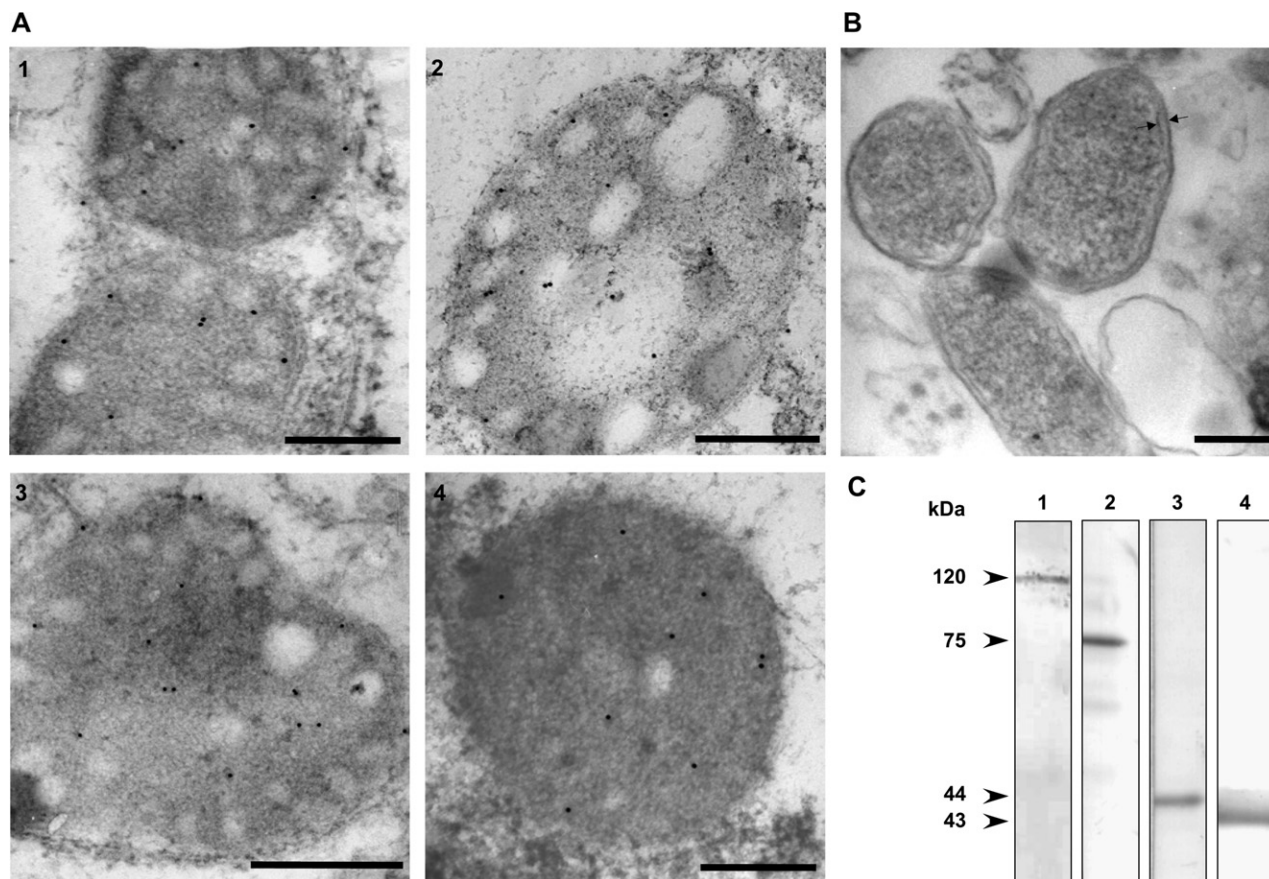


Fig. 3. *Blastocystis hominis* mitochondria-like organelles (MLOs) contain both hydrogenosomal and typical mitochondrial proteins. (A) Transmission electron microscopy (TEM) immunogold labelling of *B. hominis* sections using antibodies produced against pyruvate ferredoxin oxidoreductase (PFOR) (1), iron-hydrogenase (2), NADH:ubiquinone oxidoreductase subunit 7 (NAD7) (3) and β -succinyl-CoA synthetase β -SCS) (4). A similar immunostaining pattern of the MLO matrix is observed with the four mouse antisera. (B) Electron micrograph of the purified MLOs with a better visualization of the two closely adjacent membranes (arrows). Bars, 400 nm. (C) Immunodetection of PFOR (1), iron-hydrogenase (2), NAD7 (3) and β -SCS (4) in a protein extract from a purified-MLO fraction. *Associated methods*: partial sequences of genes coding for PFOR, iron-hydrogenase, NAD7 and β -SCS were PCR-amplified and cloned into the bacterial expression vector pET-23b (Novagen). Expression of the His-tagged recombinant proteins (PFOR amino acids 1–1223, iron-hydrogenase amino acids 157–759, NAD7 amino acids 15–230 and β -SCS amino acids 255–425) was analysed in *Escherichia coli* BL21 Star strain after isopropyl-beta-D-thiogalactopyranoside (IPTG) induction. Mice were injected i.p. with each purified recombinant protein homogenised with FCA (Sigma), and identical injections were given on days 14 and 21 with Freund incomplete adjuvant (Sigma). Sera were collected 1 week after the last injection. For immunogold staining (A), mouse antisera were used at a 1:50 dilution. Proteins from the purified MLOs, solubilised by Laemmli buffer 15 min at 100 °C, were subjected to SDS-PAGE and transferred onto polyvinylidene fluoride (PVDF) membranes (Immobilon-P; Millipore) for Western blot analysis (B). The membranes were incubated with sera (1:300 dilution) and with horseradish peroxidase-conjugated goat anti-mouse IgG, A, M (1:10,000, Sigma). Bound antibodies were revealed using NBT-BCIP (Promega).

For immunofluorescence and immunoelectron microscopy assays, polyclonal antibodies were raised against four proteins: PFOR, iron-hydrogenase, the SCS β subunit (β -SCS) and the NADH:ubiquinone oxidoreductase subunit 7 (NAD7, its gene being harboured by the circular DNA). In immunofluorescence assays, a punctuate staining was observed on whole cells with each polyclonal serum, suggesting that the four studied proteins were compartmentalised within the *B. hominis* cells (data not shown) as demonstrated for iron-hydrogenase by Stechmann et al. (2008). Immunoelectron microscopy demonstrated that these four proteins were precisely located in the MLO matrix (Fig. 3A). We observed two morphological types of MLOs, one with cristae as described previously (Zierdt et al., 1988) and one without. Both types could co-exist in a given parasite and were stained by our antibodies

and the oligonucleotide probe (Figs. 2 and 3, and data not shown).

We developed a protocol for the purification of *B. hominis* MLOs based on the procedures of Drmotá et al. (1996) and Nasirudeen and Tan (2004). Briefly, 7×10^8 parasites were resuspended in isotonic medium (0.25 M sucrose, 10 mM potassium phosphate, 1 mM EDTA, pH 7.5) and disrupted in a Potter-Elvehjem tissue homogeniser. After lysate centrifugation, the MLO-rich fraction was resuspended in the isotonic medium and then mixed with an equal volume of 90% Percoll (GE Healthcare Bio-Sciences) in DHE buffer (2 mM dithiothreitol, 30 mM HEPES, 1 mM EDTA, pH 7.5). The purified MLOs were then collected by centrifugation at 80,000g for 2 h 30 min and resuspended in DHE buffer. Western blot analyses were performed on proteins extracted from the purified MLO

fraction (Fig. 3B). Sera against β -SCS, NAD7, iron-hydrogenase and PFOR each reacted with a single band at 43, 44, 75, and 120 kDa, respectively (Fig. 3C). This was consistent with the predicted molecular sizes, and validated the functionality of *B. hominis* MLO translational machinery.

Our data indicate that *B. hominis* MLOs are beyond doubt mitochondria even though they lack the cytochromes and cytochrome oxidase (Zierdt, 1986). Genes coding ssrRNAs and lsrRNAs, tRNAs, and electron transport chain (ETC) complex I subunits are clearly typically mitochondrial, as strengthened by previous phylogenetic analyses showing their clustering with mitochondrial clades linked to α -proteobacteria (Stechmann et al., 2008). Thus, the two Stramenopiles *B. hominis* and *Phytophthora* spp. should share a common ancestor, and their organelles should probably originate from an endosymbiotic α -proteobacteria. Losses of both the ETC components and the ability to produce energy by oxidative phosphorylation are probably due to an adaptive evolution to anaerobic or microaerophilic environments. *B. hominis* organelles have the mitochondrial ETC complexes I and II and these may produce the proton gradient previously shown by biochemical studies (Nasirudeen and Tan, 2004; Stechmann et al., 2008). In parasitic helminth anaerobic mitochondria, malate is reduced to succinate by two reactions, one functioning in the reverse direction of the tricarboxylic acid (TCA) cycle. These steps are essential for the redox balance providing cofactor regeneration. In this case fumarate is the terminal electron acceptor of the ETC allowing ATP-formation (Tielens et al., 2002). This could be the same for *B. hominis*, according to preliminary nuclear magnetic resonance (NMR) studies that argue for succinate as the main produced metabolite (data not shown).

Strikingly, we found specific hydrogenosomal enzymes in the MLOs (PFOR and iron-hydrogenase). According to genome data, *B. hominis* possesses only one hydrogenase and this is rather unusual as it is fused with a flavodoxin domain. A similar hydrogenase was recently described in the *Trichomonas vaginalis* genome draft (Carlton et al., 2007). Nonetheless we could not obtain evidence for hydrogen production nor could we detect any D2/H⁺ exchange (signature of hydrogenase catalytic site, Jouanneau et al., 1980) in our *B. hominis* culture conditions (data not shown), calling into question the role of this enzyme. We thus incline towards the hypothesis that the MLOs are anaerobic mitochondria. Our data thus reinforce the theory of α -proteobacteria endosymbiosis for double-membrane surrounded-organelles in *B. hominis*. The discussion on MLO metabolism and phylogenetic issues will be addressed more in detail in a future paper on the complete nuclear genome.

Acknowledgments

M.R. was supported by a grant from the Ministère de l'Éducation Nationale de l'Enseignement Supérieur et de

la Recherche (MENESR). We are grateful to Laurent Cournac (CEA Cadarache IBEB/SBVME, Laboratoire de Bioénergétique et Biotechnologie des Bactéries et Microalgues, France) for experiments concerning hydrogen production in *B. hominis* cultures. We also thank Muriel Mazet for her precious advice for MLO purification optimisation and Damien Brosson for electron micrograph (Fig. 2A). Finally, we wish to thank Charles White (UMR CNRS 6247 Génétique, Reproduction et Développement, Université Blaise Pascal, France) for helpful reading of the paper.

References

- Barberà, M.J., Ruiz-Trillo, I., Leigh, J., Hug, L.A., Roger, A.J., 2007. The diversity of mitochondrion-related organelles amongst eukaryotic microbes. In: Martin, W.F., Müller, M. (Eds.), *Origin of Mitochondria and Hydrogenosomes*. Springer-Verlag, Berlin, pp. 239–275.
- Carlton, J.M., Hirt, R.P., Silva, J.C., Delcher, A.L., Schatz, M., Zhao, Q., Wortman, J.R., Bidwell, S.L., Alsmark, U.C., Besteiro, S., Sicheritz-Ponten, T., Noel, C.J., Dacks, J.B., Foster, P.G., Simillion, C., Van de Peer, Y., Miranda-Saavedra, D., Barton, G.J., Westrop, G.D., Muller, S., Dessi, D., Fiori, P.L., Ren, Q., Paulsen, I., Zhang, H., Bastida-Corcuer, F.D., Simoes-Barbosa, A., Brown, M.T., Hayes, R.D., Mukherjee, M., Okumura, C.Y., Schneider, R., Smith, A.J., Vanacova, S., Villalvazo, M., Haas, B.J., Perte, M., Feldblyum, T.V., Utterback, T.R., Shu, C.L., Osoegawa, K., de Jong, P.J., Hrdy, I., Horvathova, L., Zubacova, Z., Dolezal, P., Malik, S.B., Logsdon Jr., J.M., Henze, K., Gupta, A., Wang, C.C., Dunne, R.L., Upcroft, J.A., Upcroft, P., White, O., Salzberg, S.L., Tang, P., Chiu, C.H., Lee, Y.S., Embley, T.M., Coombs, G.H., Mottram, J.C., Tachezy, J., Fraser-Liggett, C.M., Johnson, P.J., 2007. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. *Science* 315, 207–212.
- Drmota, T., Proost, P., Van Ranst, M., Weyda, F., Kulda, J., Tachezy, J., 1996. Iron-ascorbate cleavable malic enzyme from hydrogenosomes of *Trichomonas vaginalis*: purification and characterization. *Mol. Biochem. Parasitol.* 83, 221–234.
- Embley, T.M., Martin, W., 2006. Eukaryotic evolution, changes and challenges. *Nature* 440, 623–630.
- Hackstein, J.H., Tjaden, J., Huynen, M., 2006. Mitochondria, hydrogenosomes and mitosomes: products of evolutionary tinkering! *Curr. Genet.* 50, 225–245.
- Ho, L.C., Singh, M., Suresh, G., Ng, G.C., Yap, E.H., 1993. Axenic culture of *Blastocystis hominis* in Iscove's modified Dulbecco's medium. *Parasitol. Res.* 79, 614–616.
- Jouanneau, Y., Kelley, B.C., Berlier, Y., Lespinat, P.A., Vignais, P.M., 1980. Continuous monitoring, by mass spectrometry, of H₂ production and recycling in *Rhodospseudomonas capsulata*. *J. Bacteriol.* 143, 628–636.
- Martin, F.N., Bensasson, D., Tyler, B.M., Boore, J.L., 2007. Mitochondrial genome sequences and comparative genomics of *Phytophthora ramorum* and *P. sojae*. *Curr. Genet.* 51, 285–296.
- Mazet, M., Diogon, M., Alderete, J.F., Vivarès, C.P., Delbac, F., 2008. First molecular characterisation of hydrogenosomes in the protozoan parasite *Histomonas meleagridis*. *Int. J. Parasitol.* 38, 177–190.
- Nasirudeen, A.M., Tan, K.S., 2004. Isolation and characterization of the mitochondrion-like organelle from *Blastocystis hominis*. *J. Microbiol. Methods* 58, 101–109.
- Patterson, D.J., 1999. The diversity of eukaryotes. *Am. Nat.* 154, S96–S124.
- Stark, D., van Hal, S., Marriott, D., Ellis, J., Harkness, J., 2007. Irritable bowel syndrome: a review on the role of intestinal protozoa and the importance of their detection and diagnosis. *Int. J. Parasitol.* 37, 11–20.
- Stechmann, A., Hamblin, K., Perez-Brocal, V., Gaston, D., Richmond, G.S., van der Giezen, M., Clark, C.G., Roger, A.J., 2008. Organelles in

- Blastocystis* that blur the distinction between mitochondria and hydrogenosomes. *Curr. Biol.* 18, 580–585.
- Stensvold, C.R., Suresh, G.K., Tan, K.S., Thompson, R.C., Traub, R.J., Viscogliosi, E., Yoshikawa, H., Clark, C.G., 2007. Terminology for *Blastocystis* subtypes – a consensus. *Trends Parasitol.* 23, 93–96.
- Taupin, V., Méténier, G., Delbac, F., Vivarès, C.P., Prensier, G., 2006. Expression of two cell wall proteins during the intracellular development of *Encephalitozoon cuniculi*: an immunocytochemical and in situ hybridization study with ultrathin frozen sections. *Parasitology* 132, 815–825.
- Tielens, A.G., Rotte, C., van Hellemond, J.J., Martin, W., 2002. Mitochondria as we don't know them. *Trends Biochem. Sci.* 27, 564–572.
- Vogel, R.O., Smeitink, J.A., Nijtmans, L.G., 2007. Human mitochondrial complex I assembly: a dynamic and versatile process. *Biochim. Biophys. Acta* 1767, 1215–1227.
- Zierdt, C.H., 1986. Cytochrome-free mitochondria of an anaerobic protozoan – *Blastocystis hominis*. *J. Protozool.* 33, 67–69.
- Zierdt, C.H., Donnelly, C.T., Muller, J., Constantopoulos, G., 1988. Biochemical and ultrastructural study of *Blastocystis hominis*. *J. Clin. Microbiol.* 26, 965–970.

II. Prédiction du sécrétome de *Blastocystis* sp. ST7 et identification de potentiels facteurs de virulence

Afin d'évaluer le caractère potentiellement pathogène de ce microorganisme nous avons effectué une recherche des potentiels facteurs de virulence du parasite. *Blastocystis* sp. ST7 possède ainsi dans son génome un arsenal de gènes codant des protéines connues pour être des facteurs de virulence chez d'autres organismes. La prédiction du sécrétome de *Blastocystis* sp. ST7 afin d'identifier les protéines à l'interface entre l'hôte et le parasite a révélé des protéines pouvant être reliées à des phénomènes de pathogénie (protéases, glycosidases, hydrolases, inhibiteurs de protéases) (Denoeud et al., 2011). Ces différentes protéines candidates pourraient participer à l'attaque du mucus et des cellules épithéliales au niveau intestinal ou à la modulation de la réponse immune. Ces effets pourraient expliquer les symptômes gastro-intestinaux aspécifiques rencontrés chez certains porteurs du parasite. Ces hypothèses nécessitent toutefois d'être démontrées expérimentalement pour apporter des données concrètes sur la physiopathologie et les interactions que peut présenter ce parasite avec son hôte.

Nous avons également mis en évidence des gènes codant des polyketide synthases (PKS) de type I et des NRPS (non-ribosomal peptide synthase) qui synthétisent respectivement des polyketides et des peptides non ribosomiques. Les polyketides sont des métabolites secondaires possédant des propriétés pharmacologiques intéressantes comme des antibiotiques ou des immunomodulateurs. Les peptides non ribosomiques, tel que les macrolides, peuvent avoir une activité antibactérienne. La production d'antibactériens pourrait participer à la dysbiose souvent observée dans les pathologies intestinales, comme pour le syndrome du côlon irritable (IBS, irritable bowel syndrome) (Lee et Bak, 2011). C'est dans ce contexte que nous avons été sollicités par la revue Plos Pathogens pour écrire une synthèse sur l'association IBS/*Blastocystis* spp.

Pearls

New Insights into *Blastocystis* spp.: A Potential Link with Irritable Bowel Syndrome

Philippe Poirier^{1,2,3,4}, Ivan Wawrzyniak^{1,2}, Christian P. Vivarès^{1,2}, Frédéric Delbac^{1,2}, Hicham El Alaoui^{1,2*}

1 Clermont Université, Université Blaise Pascal, Laboratoire Microorganismes, Génome et Environnement, BP 10448, Clermont-Ferrand, France, **2** CNRS, UMR 6023, LMGE, Aubiere, France, **3** Clermont Université, Université d'Auvergne, JE 2526, Evolution des bactéries pathogènes et susceptibilité de l'hôte, BP 10448, Clermont-Ferrand, France, **4** CHU Clermont-Ferrand, Service Parasitologie Mycologie, Clermont-Ferrand, France

What Are *Blastocystis* spp.?

Blastocystis spp. belong to the phylum Stramenopila, a complex and heterogeneous evolutionary assemblage of heterotrophic and photosynthetic protozoa [1]. Interestingly, this is the only stramenopile living in the lower digestive tract of humans, and it also lives in other mammals, birds, reptiles, amphibians, and insects [1]. Even though isolates were reported to be morphologically indistinguishable, an extensive genetic variation among isolates from both humans and animals has been observed. Thirteen subtypes (ST1–ST13), with the first nine being found in humans, have been identified based on genes coding for the small-subunit ribosomal RNA [2]. Preferential repartition of STs exists among animals that appear to constitute the main reservoir for environmental dissemination and human contamination [1].

Four forms of *Blastocystis* spp. (vacuolar, granular, amoeboid, and cyst) were described in stools and/or in vitro cultures [1]. Studies in animals demonstrated that the water- and environmentally resistant infective cyst undoubtedly represents the transmissible stage of this parasite [1]. *Blastocystis* spp. prevalence in humans often exceeds 5% in industrialized countries and can reach as high as 76% in developing countries [1,3]. However, prevalence data are largely dependent on the methods used for detection, quantitative PCR being the most sensitive method, meaning that infections by *Blastocystis* spp. are likely underestimated [4].

Lately, *Blastocystis* spp. have been included in the water sanitation and health programs of the World Health Organization [5]. Increasing interest of scientific and medical communities for *Blastocystis* spp. was coupled with new data about epidemiology, pathogenicity, and, more recently, the first whole genome of a human isolate. Accumulating in vivo, in vitro, and in silico data has enabled researchers to assess the potential impact of *Blastocystis* spp. in human health.

Are *Blastocystis* spp. Pathogens?

In vivo endoscopy and biopsy analyses in symptomatic patients indicated that *Blastocystis* spp. do not invade the colonic mucosa, but lead to disturbances on the barrier function and permeability [1,6]. Experiments on immunocompetent BALB/c mice revealed intense inflammatory-cell infiltration in the mucosa of some specimens, but not in all mice, suggesting that some host factors could be involved [1]. Subsequently, the infectivity of human isolates obtained from both asymptomatic and symptomatic patients on rats was assessed by Hussein et al. [7]. Interestingly, the moderate and severe degrees of pathological changes were only found in rats infected by isolates from symptomatic patients, and differences in severity were observed among the different STs of *Blastocystis*, suggesting the existence of some more virulent strains.

To understand cellular mechanisms, in vitro experiments were performed to investigate the cytopathic effects of *Blastocystis* spp. on mammalian cell cultures (Figure 1). A first study showed in the rat epithelial cell line IEC-6 that *Blastocystis* ST4 can induce apoptosis in a contact-independent manner, increasing epithelial permeability [8]. The pro-inflammatory effect of *Blastocystis* ST1 culture filtrates was demonstrated on HT-29 and T-84 human colonic epithelial cells with production of interleukin 8 (IL-8) and granulocyte-macrophage colony stimulating factor (GM-CSF) [1]. Cysteine proteases of *Blastocystis* ST4 were shown to induce IL-8 production via an NF-κB pathway [9]. Proteases released in culture supernatants of both *Blastocystis* ST4 and ST7 were also shown to be able to cleave human-secreted immunoglobulin A (IgA) and then modulate the immune response of the host [1]. A surface-located cysteine protease was recently shown to be involved in a pro-survival role in *Blastocystis* ST7 and may activate other proteases [10]. Nevertheless, in vitro studies are limited by the lack of tools to study *Blastocystis* spp. Indeed, few strains of *Blastocystis* spp. are available in axenic cultures, and growth rates in culture are fluctuating and quite low. In addition, growth of this parasite is realized in anaerobic chambers that limit the possibility of long-term exposure in cellular models.

Is There a Link between *Blastocystis* spp. and IBS?

Some studies have suggested an association of *Blastocystis* spp. with acute or chronic digestive disorders such as irritable bowel syndrome (IBS) [1]. IBS is a functional gastrointestinal disorder characterized by abdominal discomfort and/or pain associated with changes in bowel habits, affecting 5%–24% of people in industrialized countries with impairment on quality of life [11]. In 1997, Hussain et al. highlighted that sera from IBS patients were characterized by higher IgG antibody levels to *Blastocystis* spp. when compared to healthy populations [12]. However, the first relevant epidemiological report about a possible link between

Citation: Poirier P, Wawrzyniak I, Vivarès CP, Delbac F, El Alaoui H (2012) New Insights into *Blastocystis* spp.: A Potential Link with Irritable Bowel Syndrome. PLoS Pathog 8(3): e1002545. doi:10.1371/journal.ppat.1002545

Editor: Laura J. Knoll, University of Wisconsin Medical School, United States of America

Published: March 15, 2012

Copyright: © 2012 Poirier et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors received no specific funding for this article.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: hicham.el_alaoui@univ-bpclermont.fr

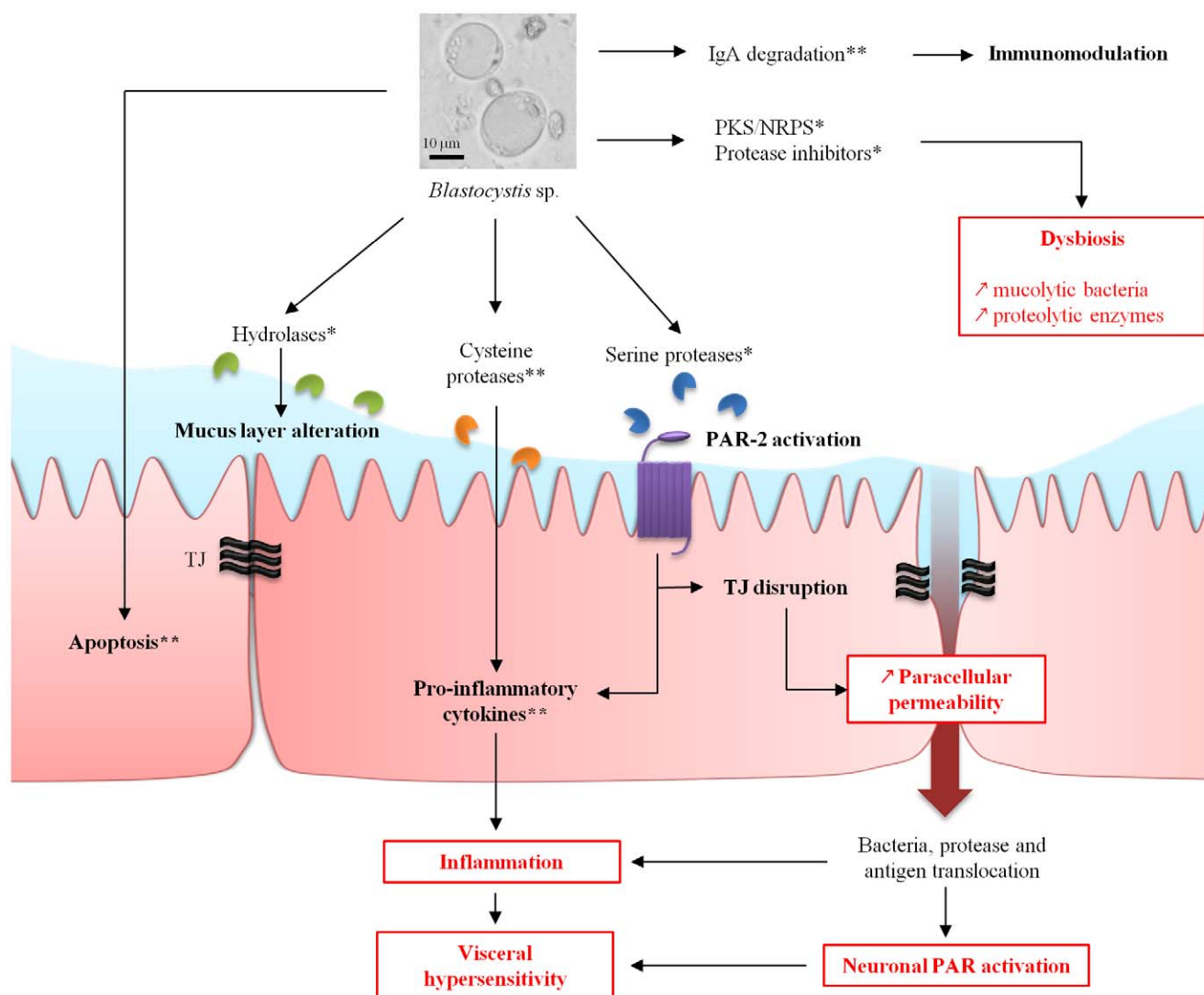


Figure 1. Hypothetical model of pathogeny for *Blastocystis* spp. from genomic (*) and experimental () data.** Potential link to IBS pathophysiology mechanisms (in red). IgA, immunoglobulin A; NRPS, non-ribosomal polyketide synthase; PAR, protease-activated receptor; PKS, polyketide synthase.
doi:10.1371/journal.ppat.1002545.g001

Blastocystis spp. and IBS was provided two years later by Giacometti et al. [13]. When comparing the prevalence of *Blastocystis* spp. in individuals with gastrointestinal symptoms and classified as affected or not by IBS, the authors found that it was significantly present in IBS patients. Other more recent studies also argued for a higher prevalence of *Blastocystis* spp. among IBS patients (Table 1) compared to healthy populations or to patients suffering from other gastrointestinal disorders [3,13–17]. However, three studies failed to demonstrate an association between *Blastocystis* spp. and IBS [18–20]. Possible explanations could be the small IBS cohort studied or the parasitological diagnostic methods used. Two of these three studies were from Thailand and one from Mexico, whereas studies arguing for a link between *Blastocystis* spp. and IBS were from the Middle East (Table 1) and Europe. IBS is a functional disorder of multifactorial origin, and some genetic, environmental, and microbiological factors could also explain this discrepancy. Nevertheless, overall studies missed an opportunity to provide a therapeutic trial with a follow up of IBS symptoms.

Recent studies suggest that visceral pain associated with IBS could be explained by alterations of the epithelial barrier, resulting in bowel motility and sensitivity disorders (Figure 1). Indeed, in vitro studies on colonic biopsies from patients with IBS showed an increase of paracellular permeability associated with perturbations of tight junctions (TJs) [21]. On the other hand, it is now well recognized that there is a low-grade inflammation of the mucosa in IBS patients [22]. Then, protease-activated receptor type 2 (PAR-2) was proposed to be involved in both an increase of permeability and low-grade inflammation [23]. PAR-2 are activated by serine-proteases that cleave the N-terminal domain of the receptor. Then, the released peptide may act as a ligand and turn on the receptor to enhance TJ opening and trigger inflammation. Increase of paracellular permeability allows diffusion of both antigens and bacteria to sub-mucosa, participating in inflammation. PARs are also present at the surface of intestinal neurons. The activation of some members of the PAR family could contribute to abdominal pain [24]. Studies also showed that stools from IBS patients present higher proteolytic activity than healthy

Table 1. Summary of studies investigating the association between *Blastocystis* spp. and IBS.

Studies	Country	Diagnostic Method	IBS Patients		Control Group	
			Prevalence	n (%)	Criteria ^b	Prevalence n (%) Population
Giacometti et al. [13]	Italy	Trichrome stain	15/81	(18.5) ^a	Rome	23/307 (7.5) GD
Yakoob et al. [17]	Pakistan	Culture	44/95	(46) ^a	Rome II	4/55 (7) GD
Tungtrongchitr et al. [20]	Thailand	Culture	8/59	(13.6)	Rome II	3/25 (12) Healthy
Yakoob et al. [15]	Pakistan	Culture	95/158	(60) ^a	Rome III	38/157 (24) Healthy, GD
Surangsri et al. [19]	Thailand	Culture	11/66	(16.7)	Rome II	6/60 (10) GD
Yakoob et al. [16]	Pakistan	Culture	90/171	(53) ^a	Rome III	25/159 (16) Healthy
Dogruman-AI et al. [3]	Turkey	Lugol stain	8/21	(38) ^a	Rome III	5/43 (11.6) Healthy
Ramirez-Miranda et al. [18]	Mexico	Flotation method	n.a./115	(15.7)	Rome III	n.a./209 (12) GD
Jimenez-Gonzalez et al. [14]	Mexico	Flotation method	14/45	(31.1) ^a	Rome III	6/45 (13.3) GD

^aStudies concluding to an association between *Blastocystis* spp. and IBS.

^bRome criteria used to diagnose IBS.

GD, gastrointestinal disorders; n.a., not available.

doi:10.1371/journal.ppat.1002545.t001

controls [25]. Experimental data have shown a protease activity of supernatants from axenic cultures of both *Blastocystis* ST4 and ST7 [1]. This is supported by the prediction of 22 secreted proteases from genomic data of *Blastocystis* ST7 (detailed below) [26]. These experimental and genomic data suggest a possible involvement of parasite proteases in gastrointestinal disturbances. Thus, proteases from bacteria or *Blastocystis* spp., such as metalloproteases, cysteine, or serine proteases, could play a key role in IBS genesis [24]. This perturbation could be linked to a modification of lumen microbiota in IBS patients compared to healthy patients [27]. Then, dysbiosis may take part in low-grade inflammation of the mucosa and IBS symptoms.

How Can Genomic Data Support the *Blastocystis* spp./IBS Association?

The whole genome of a *Blastocystis* ST7 isolate has been sequenced [26]. Interestingly, candidate proteins potentially involved in the pathogenicity of *Blastocystis* spp. were identified by in silico analyses of the predicted proteome and secretome (Figure 1).

Blastocystis ST7 likely uses hydrolases to attack host tissues for its nutrient supply. Fucosidase, hexosaminidase, and β -galactosidase were identified in the *Blastocystis* ST7 predicted secretome. *Blastocystis* ST7 may participate in this process by degrading host glycoproteins, especially those that constitute the mucus [26]. In addition, cysteine proteases could also degrade mucins [24]. Thus, these enzymes may allow *Blastocystis* spp. to use mucus as a carbohydrate and protein source and enable them to survive within the intestinal environment by creating their own micro-environment. Interestingly, 22 proteases, including 20 cysteine proteases, one serine protease, and one aspartic protease, were predicted to be secreted [26]. The impairment of mucus may be the initial step to inflammatory and allergic disturbance caused by chronic exposure to luminal antigens. Then, proteases from *Blastocystis* spp. and/or gut bacteria can also target receptors at the intestinal cell surface. These proteolytic enzymes are known to be involved in paracellular permeability, inflammation, and hypersensitivity [24]. The *Blastocystis* ST7 serine protease could therefore have the ability to target PAR-2 (Figure 1), inducing inflammation and TJ disruption as frequently seen in IBS. Once the TJs are opened, luminal proteases can have access to submucosal ganglia,

activate PARs on enteric neurons (Figure 1), and be responsible for hypersensitivity in IBS patients [24]. Penetration of luminal bacteria or antigens could participate in the establishment of a chronic low-grade inflammation in submucosa by stimulation of innate immunity.

Blastocystis ST7-secreted glycosyltransferases could also participate in TJ disruption in IBS. This can be illustrated by the lymphostatin (toxin) of *Citrobacter rodentium*, which possesses a glycosyltransferase activity that primarily influences localization of ZO-1 and occludin at intestinal cell TJs, compromising epithelial barrier function [24]. This is also the case for *Clostridium* spp., which exhibit toxins with glycosyltransferase activities [24]. These proteins act by inactivating Rho proteins that are known to be important in maintaining TJs.

Concerning the intestinal protease balance that regulates gut functioning, protease inhibitors released by enteric pathogens or parasites can modulate the activity of host proteases and disturb intestinal homeostasis [28]. Genes coding for protease inhibitors are also present in the *Blastocystis* ST7 genome, and some are predicted to be secreted, including cystatin, type 1-protease inhibitor, and endopeptidase inhibitor-like protein [26]. Moreover, dysbiosis was shown to occur during IBS [27]. Genomic data revealed that a polyketide synthase (PKS) and a non-ribosomal polyketide synthase (NRPS) are present in *Blastocystis* ST7. These enzymes are known to produce non-ribosomal peptides and polyketides with various biological properties, such as antibiotics or immunomodulatory molecules that could participate in dysbiosis and inflammation [29].

Finally, we can assume that *Blastocystis* spp.-secreted proteins have the potential to modulate host defenses and to facilitate nutrient acquisition and parasite colonization. Moreover, *Blastocystis* spp. would be able to alter integrity of gut epithelia and probably participate in dysbiosis.

What Is Needed to Improve the Knowledge on *Blastocystis* spp. Biology?

A consortium including the most relevant researchers on this topic is being set up. The first aim is to develop efficient standardized tools to study *Blastocystis* spp., including the improvement of cultural methods (in particular, axenization protocols of the other STs) and development of new molecular

markers for epidemiological studies. The role of STs in pathogenicity needs to be elucidated, notably by sequencing the whole genome of other STs. Comparative genomic analyses will then be useful for both identification and validation of in silico-predicted virulence factors. Animal models are also needed to go further in the understanding of the role of *Blastocystis* spp. in gut dysfunctions. Another aim is to develop some reverse genetic tools for functional characterization of genes of interest. Considering the clinical implication of *Blastocystis* spp. in IBS, none of the published

studies provided complete clinical data about IBS classification and *Blastocystis* STs. Thus, well-designed multi-centric studies are required, and studies on *Blastocystis* spp. susceptibility to treatment have to be used to suggest guidelines for the eradication of this infection [30]. The role of *Blastocystis* spp. in dysbiosis and the exact interactions with bacteria also have to be highlighted. This approach seems more relevant because of the increasing interest in microbiota disturbances in the genesis of various gastrointestinal dysfunctions.

References

1. Tan KS (2008) New insights on classification, identification, and clinical relevance of *Blastocystis* spp. Clin Microbiol Rev 21: 639–665.
2. Stensvold CR, Suresh GK, Tan KS, Thompson RC, Traub RJ, et al. (2007) Terminology for *Blastocystis* subtypes: a consensus. Trends Parasitol 23: 93–96.
3. Dogruman-Al F, Simsek Z, Boorum K, Ekici E, Sahin M, et al. (2010) Comparison of methods for detection of *Blastocystis* infection in routinely submitted stool samples, and also in IBS/IBD Patients in Ankara, Turkey. PLoS ONE 5: e15484. doi:10.1371/journal.pone.0015484.
4. Poirier P, Wawrzyniak I, Albert A, El Alaoui H, Delbac F, et al. (2011) Development and evaluation of a real-time PCR assay for detection and quantification of *blastocystis* parasites in human stool samples: prospective study of patients with hematological malignancies. J Clin Microbiol 49: 975–983.
5. WHO (2008) Guidelines for drinking-water quality. Third edition. Incorporating first and second addenda. Geneva: WHO. 514 p.
6. Dagci H, Ustun S, Taner MS, Ersoz G, Karacasu F, et al. (2002) Protozoan infections and intestinal permeability. Acta Trop 81: 1–5.
7. Hussein EM, Hussein AM, Eida MM, Atwa MM (2008) Pathophysiological variability of different genotypes of human *Blastocystis hominis* Egyptian isolates in experimentally infected rats. Parasitol Res 102: 853–860.
8. Puthia MK, Sio SW, Lu J, Tan KS (2006) *Blastocystis ratti* induces contact-independent apoptosis, F-actin rearrangement, and barrier function disruption in IEC-6 cells. Infect Immun 74: 4114–4123.
9. Puthia MK, Lu J, Tan KS (2008) *Blastocystis ratti* contains cysteine proteases that mediate interleukin-8 response from human intestinal epithelial cells in an NF-kappaB-dependent manner. Eukaryot Cell 7: 435–443.
10. Wu B, Yin J, Texier C, Roussel M, Tan KS (2010) *Blastocystis* legumain is localized on the cell surface, and specific inhibition of its activity implicates a pro-survival role for the enzyme. J Biol Chem 285: 1790–1798.
11. Longstreth GF, Thompson WG, Chey WD, Houghton LA, Mearin F, et al. (2006) Functional bowel disorders. Gastroenterology 130: 1480–1491.
12. Hussain R, Jaferi W, Zuberi S, Baqai R, Abrar N, et al. (1997) Significantly increased IgG2 subclass antibody levels to *Blastocystis hominis* in patients with irritable bowel syndrome. Am J Trop Med Hyg 56: 301–306.
13. Giacometti A, Cirioni O, Fiorentini A, Fortuna M, Scalise G (1999) Irritable bowel syndrome in patients with *Blastocystis hominis* infection. Eur J Clin Microbiol Infect Dis 18: 436–439.
14. Jimenez-Gonzalez DE, Martinez-Flores WA, Reyes-Gordillo J, Ramirez-Miranda ME, Arroyo-Escalante S, et al. (2011) *Blastocystis* infection is associated with irritable bowel syndrome in a Mexican patient population. Parasitol ResEpub ahead of print 26 August 2011. doi:10.1007/s00436-011-2626-7.
15. Yakoob J, Jafri W, Beg MA, Abbas Z, Naz S, et al. (2010) Irritable bowel syndrome: is it associated with genotypes of *Blastocystis hominis*. Parasitol Res 106: 1033–1038.
16. Yakoob J, Jafri W, Beg MA, Abbas Z, Naz S, et al. (2010) *Blastocystis hominis* and *Dientamoeba fragilis* in patients fulfilling irritable bowel syndrome criteria. Parasitol Res 107: 679–684.
17. Yakoob J, Jafri W, Jafri N, Khan R, Islam M, et al. (2004) Irritable bowel syndrome: in search of an etiology: role of *Blastocystis hominis*. Am J Trop Med Hyg 70: 383–385.
18. Ramirez-Miranda ME, Hernandez-Castellanos R, Lopez-Escamilla E, Moncada D, Rodriguez-Magallan A, et al. (2010) Parasites in Mexican patients with irritable bowel syndrome: a case-control study. Parasit Vectors 3: 96.
19. Surangsriat S, Thamrongwittawatpong L, Piyaniran W, Naaglor T, Khoprasert C, et al. (2010) Assessment of the association between *Blastocystis* infection and irritable bowel syndrome. J Med Assoc Thai 93 Suppl 6: S119–S124.
20. Tungtrongchitr A, Manatsathit S, Kositchaiwat C, Ongrotchanakun J, Munkong N, et al. (2004) *Blastocystis hominis* infection in irritable bowel syndrome patients. Southeast Asian J Trop Med Public Health 35: 705–710.
21. Piche T, Barbara G, Aubert P, Bruley des Varannes S, Dainese R, et al. (2009) Impaired intestinal barrier integrity in the colon of patients with irritable bowel syndrome: involvement of soluble mediators. Gut 58: 196–201.
22. Chadwick VS, Chen W, Shu D, Paulus B, Bethwaite P, et al. (2002) Activation of the mucosal immune system in irritable bowel syndrome. Gastroenterology 122: 1778–1783.
23. Bueno L, Fioramonti J (2008) Protease-activated receptor 2 and gut permeability: a review. Neurogastroenterol Motil 20: 580–587.
24. Steck N, Mueller K, Schemann M, Haller D (2011) Bacterial proteases in IBD and IBS. GutEpub ahead of print 7 September 2011. doi:10.1136/gutjnl-2011-300775.
25. Gece K, Roka R, Ferrier L, Leveque M, Eutamene H, et al. (2008) Increased faecal serine protease activity in diarrhoeic IBS patients: a colonic luminal factor impairing colonic permeability and sensitivity. Gut 57: 591–599.
26. Denoeud F, Roussel M, Noel B, Wawrzyniak I, Da Silva C, et al. (2011) Genome sequence of the stramenopile *Blastocystis*, a human anaerobic parasite. Genome Biol 12: R29.
27. Lee BJ, Bak YT (2011) Irritable bowel syndrome, gut microbiota and probiotics. J Neurogastroenterol Motil 17: 252–266.
28. Rhoads ML, Fetterer RH, Hill DE, Urban JF, Jr. (2000) *Trichuris suis*: a secretory chymotrypsin/elastase inhibitor with potential as an immunomodulator. Exp Parasitol 95: 36–44.
29. Schwarzer D, Marahiel MA (2001) Multimodular biocatalysts for natural product assembly. Naturwissenschaften 88: 93–101.
30. Coyle CM, Varughese J, Weiss LM, Tanowitz HB (2012) *Blastocystis*: to treat or not to treat. Clin Infect Dis 54: 105–110.

III. Caractérisation de deux cystéines protéases sécrétées par *Blastocystis* sp ST7.

Chez les protozoaires parasites, les protéases sont connues pour jouer des rôles clés dans la biologie et la pathogénie, notamment dans la cytoadhésion, l'invasion cellulaire ou encore dans le dékystement (McKerrow et al., 1993). Chez *Blastocystis* sp., des activités protéasiques ont été rapportées à partir de lysats cellulaires (Sio et al., 2006; Mirza et Tan, 2009) et dans les surnageants de culture (Puthia et al., 2005). Ces enzymes sont capables de dégrader les IgA humaines (Puthia et al., 2005), de moduler la réponse immunitaire (Long et al., 2001; Puthia et al., 2008) ou encore d'augmenter la perméabilité de cellules épithéliales intestinales (Puthia et al., 2006). L'analyse du génome a permis de mettre en évidence 22 protéases potentiellement sécrétées. Ces protéases sécrétées sont réparties en 20 protéases à cystéine, 1 protéase à sérine et 1 protéase à aspartate. Parmi les protéases à cystéine on retrouve 9 légumaines, 3 cathepsines L, 3 cathepsines C, 3 cathepsines B, 1 caricaïne et 1 indéterminée.

Afin de confirmer ces analyses *in silico* concernant les protéases sécrétées, nous avons caractérisé les activités protéolytiques présentes dans les surnageants de culture. Ce travail est en cours de révision pour publication dans la revue *Parasitology International*. Nous avons pu confirmer la présence dans les produits de sécrétion de *Blastocystis* sp. ST7 de deux protéases prédites par l'analyse bio-informatique. Ces deux enzymes constituent de sérieux candidats dans la physiopathologie de ce parasite. En effet, ce type d'enzymes est connu pour dégrader les mucines présentes à la surface de l'épithélium intestinal. Ainsi, ces protéines permettent probablement à *Blastocystis* spp. d'utiliser le mucus comme source de protéines et de sucres et ainsi de persister dans cet environnement. La dégradation du mucus pourrait permettre à certaines protéines de *Blastocystis* sp. ST7 ou à des bactéries présentes dans l'intestin d'accéder aux cellules épithéliales provoquant ainsi leur perturbation. D'autre part, dans certains modèles de pathogènes intestinaux, ces enzymes protéolytiques sont connues pour être impliquées

dans l'augmentation de la perméabilité cellulaire, l'inflammation et l'hypersensibilité (McKerrow et al., 1993; Sajid et McKerrow, 2002). De tels symptômes sont également observés dans les pathologies intestinales comme l'IBS. Les deux protéases de *Blastocystis* sp. ST7 que nous avons identifiées pourraient donc être des effecteurs impliqués dans l'établissement de l'IBS.



Contents lists available at SciVerse ScienceDirect

Parasitology International

journal homepage: www.elsevier.com/locate/parint

Characterization of two cysteine proteases secreted by *Blastocystis* ST7, a human intestinal parasite

Ivan Wawrzyniak^{a,b}, Catherine Texier^{a,b}, Philippe Poirier^{a,b}, Eric Viscogliosi^{c,d,e,f}, Kevin S.W. Tan^g, Frédéric Delbac^{a,b}, Hicham El Alaoui^{a,b,*}

^a Clermont Université, Université Blaise Pascal, Laboratoire Microorganismes: Génome et Environnement, BP 10448, F-63000 Clermont-Ferrand, France

^b CNRS, UMR 6023, LMGE, F-63177 Aubière, France

^c University Lille-Nord de France, F-59000 Lille Cedex, France

^d Center for Infection and Immunity of Lille, Institut Pasteur de Lille, F-59019 Lille Cedex, France

^e Inserm U1019, F-59000 Lille Cedex, France

^f CNRS UMR 8402, F-59021 Lille Cedex, France

^g Laboratory of Molecular and Cellular Parasitology, Department of Microbiology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, 5 Science Drive 2, 117597 Singapore

ARTICLE INFO

Article history:

Received 24 August 2011

Received in revised form 17 February 2012

Accepted 20 February 2012

Available online xxxx

Keywords:

Blastocystis

Secreted proteases

Cathepsin

Legumain

Pathogenicity

Proteomics

ABSTRACT

Blastocystis spp. are unicellular anaerobic intestinal parasites of both humans and animals and the most prevalent ones found in human stool samples. Their association with various gastrointestinal disorders raises the questions of its pathogenicity and of the molecular mechanisms involved. Since secreted proteases are well-known to be implicated in intestinal parasite virulence, we intended to determine whether *Blastocystis* spp. possess such pathogenic factors. *In silico* analysis of the *Blastocystis* subtype 7 (ST7) genome sequence highlighted 22 genes coding proteases which were predicted to be secreted. We characterized the proteolytic activities in the secretory products of *Blastocystis* ST7 using specific protease inhibitors. Two cysteine proteases, a cathepsin B and a legumain, were identified in the parasite culture supernatant by gelatin zymographic SDS-PAGE gel and MS/MS analysis. These proteases might act on intestinal cells and disturb gut function. This work provides serious molecular candidates to link *Blastocystis* spp. and intestinal disorders.

© 2012 Elsevier Ireland Ltd. All rights reserved

1. Introduction

Blastocystis spp. are unicellular anaerobic parasites belonging to the stramenopile group and are found in the intestinal tract of humans as well as of a wide range of animals such as mammals, birds and rodents [1]. Four major forms, called vacuolar, granular, amoeboid and cystic, are described for this polymorphic protozoan parasite. A life cycle with cysts being the infectious stages is hypothesized although the transitions between the different forms remain to be elucidated [1]. *Blastocystis* spp. prevalence in human ranges from 3% to 5% in developed countries to 60% in developing ones [1]. But it is believed to be largely underestimated [2], since data significantly vary according to the methods used for diagnosis. A recent prospective study using a real-time quantitative PCR approach demonstrated a 14.5%-prevalence of *Blastocystis* spp. in France, compared to a 4%-one obtained by direct light microscopy observation [3]. Thirteen different subtypes (ST1 to ST13) are defined using a partial sequence of the

small ribosomal subunit rRNA gene [1,2]. Among them, subtypes ST1 to ST9 are recovered in human fecal samples with highly variable frequencies according to geographical areas [1]. The pathogenic potential of *Blastocystis* spp. remains controversial as the parasite is described in both asymptomatic and symptomatic patients. *Blastocystis* spp.-associated symptomatology consists in gastrointestinal disorders such as abdominal pains, diarrhea, nausea, cramps or urticarial symptoms [1,2]. *Blastocystis* spp. are also suggested to be involved in the irritable bowel syndrome (IBS) [4].

Proteases are well known to play major roles in parasitic biology and in host–pathogen interactions. In enteric parasites such as *Entamoeba* spp. or *Giardia* spp., cysteine proteases are shown to be efficient virulence factors [5,6]. *Entamoeba histolytica* trophozoites secrete more cysteine proteases than the non invasive species *Entamoeba dispar* [5], supporting the role of these enzymes in pathogenesis. In particular, proteases can alter gut integrity [7] and modulate host immune system [5]. Protease activities from *Blastocystis* ST4 and ST7 cell lysates [8,9], and *Blastocystis* ST4 culture supernatants [10] were previously reported. They were shown to be able to cleave human immunoglobulins A *in vitro* [10], to modulate inflammatory IL-8 production [11] and to increase permeability of intestinal epithelial cells [12]. Although these data strongly suggest a role of proteases in *Blastocystis* spp. physiopathology, the corresponding

* Corresponding author at: UBP, UMR CNRS 6023, LMGE, 24 Avenue des Landais, BP 80026, F-63171 Aubière, France. Tel.: +33 4 73 40 74 31; fax: +33 4 73 40 76 70.

E-mail address: hicham.el_alaoui@univ-bpclermont.fr (H. El Alaoui).

proteases are not yet identified. We investigated the available genomic data of a *Blastocystis* ST7 isolate [13] to search for predicted secreted proteases. We also intended to experimentally characterize secreted proteases in *Blastocystis* ST7 culture supernatants (Bcs).

2. Materials and methods

2.1. Parasite culture

The axenic culture of the *Blastocystis* ST7 used in this study was set up by Ng and Tan [14]. This isolate was obtained from a Singaporean patient with intestinal disorders. Parasites were grown as previously described [15] in pre-reduced Iscove's Modified Dulbecco's Medium (Gibco), supplemented with 10% inactivated horse serum (PAA, Pasching, Austria) and 1% penicillin/streptomycin (Gibco). Cultures were incubated at 37 °C in anaerobic conditions using the AnaeroJar system (Oxoid, Thermo Fisher Scientific).

2.2. *Blastocystis* ST7 protein sequence analyses

Protein sequences and their functional annotation were obtained from the *Blastocystis* ST7 genome sequence analysis [13]. Secreted proteases were identified according to the presence of an amino-terminal signal sequence and to protein location as predicted by SignalP 3.0 (<http://www.cbs.dtu.dk/services/SignalP>) and pSORTII (<http://psort.hgc.jp/form2.html>) algorithms, respectively. Family classification of proteases was determined using MEROPS BLAST (<http://merops.sanger.ac.uk>). If necessary, the open reading frames delineated during genome annotation were refined, the start methionine being re-positioned according to the protein size of their orthologs. Some introns were also modified using the Genewise II software (<http://www.ebi.ac.uk/Tools/Wise2/>). For proteins identified by mass spectrometry, the consensus secondary structure predictions were performed on NPS@ Network Protein Sequence Analysis (<http://npsa-pbil.ibcp.fr>) using the DSC, GOR4, PHD and SOPMA algorithms. Conserved protein motifs and domains were searched using MotifScan (<http://hits.isb-sib.ch/cgi-bin/PFSCAN>) and Pfam (<http://pfam.sanger.ac.uk/>). Comparative analysis of the *Blastocystis* ST7 cathepsin B protein with its orthologs was performed using the Clustalw2 algorithm (<http://www.ebi.ac.uk/Tools/clustalw2/>).

2.3. Preparation of culture supernatant and parasite lysate

Three-day-old parasite cultures (15 ml) were centrifuged at 1000g for 10 min. The culture supernatant was collected, centrifuged again at 1000g for 10 min, and filtrated through 0.22 µm-membranes (Millipore). It was then applied on Amicon Ultra-15 100,000 NMWL (Millipore) at 4000g during 40 min at 4 °C to remove proteins with a molecular weight higher than 100 kDa. The culture supernatant was finally concentrated using Amicon Ultra-15 10,000 NMWL at 4000g during 40 min at 4 °C, and proteins re-suspended in 2 ml of Tris-HCl 50 mM pH 7.5, and stored at -80 °C (Bcs extract). In parallel, the parasite pellet (10⁷ cells) was lysed by 10 freeze-thaw cycles (liquid nitrogen, 37 °C) and 10 sonication pulses (10×1 min on ice) using a Vibra-cell sonicator (Sonics & Materials, Newtown, CT). After centrifugation at 16,000g for 10 min at 4 °C, the clarified cell lysate was collected and stored at -80 °C (*Blastocystis* ST7 cell lysate (Bcl) extract).

2.4. Proteolytic activity assays

Protease activities in Bcs and Bcl were determined using azocasein assay [9,11]. Briefly 100 µl of each sample was incubated with 2 mM dithiothreitol (Sigma) at 37 °C for 10 min. One hundred microliters of azocasein (Sigma) at 5 mg/ml in 100 mM Tris-HCl pH 7.6 was added and the reaction was stopped after 1 h at 37 °C by addition of 300 µl of 10% trichloroacetic acid. The mixture was allowed to stand

on ice for 30 min before centrifugation at 16,000g for 10 min at 4 °C. The collected supernatant was neutralized by adding 500 µl of 1 M NaOH. Absorbance was measured with a spectrophotometer at 440 nm (Ultrospec 2000 Pharmacia Biotech). Lysate samples boiled at 100 °C for 15 min to inactivate proteases, and parasite-free medium were used as negative controls. One hundred microliters of trypsin (2.5 mg/ml) was used as a positive control. The pH dependency of *Blastocystis* ST7 protease activities was studied in reducing conditions (2 mM DTT) using 100 µl of azocasein 5 mg/ml either in 100 mM sodium acetate pH 5.5, or in 100 mM Tris-HCl pH 7, or in 200 mM NaCl, 50 mM Tris-HCl pH 9. The following protease inhibitors (Sigma) were evaluated on Bcl and Bcs extracts, in the same conditions than described above, for their ability to inhibit proteolytic activities: pepstatin A (100 µg/ml), phenylmethylsulfonylfluoride (PMSF, 1 mM), leupeptin (100 µg/ml), and trans-epoxy-succinyl-L-leucylamido-(4-guanidino) butane (E64, 1 µM). Significance was determined by Student's *t*-test.

2.5. SDS-PAGE and gelatin zymograms

For zymogram, SDS-PAGE was performed without reducing agent and heating step. Analytic 10%-acrylamide gels supplemented with 0.2% gelatin (Sigma) were used for the resolution and the recovery of proteolytic bands. Twenty microliters of Bcl (0.6 mg/ml) or Bcs (10.4 mg/ml) samples was mixed with 4× loading buffer (3% SDS, 150 mM Tris-HCl, pH 6.8, 30% glycerol, 0.1% bromophenol blue) before migration. After electrophoresis (25 mA/gel during 2 h at 4 °C), gelatin-containing gels were first rinsed in PBS and then twice in PBS containing 2.5% Triton X-100 at room temperature for 30 min. After incubation at 37 °C overnight in a development buffer (100 mM sodium acetate, 2 mM DTT, pH 5.5), they were stained with Coomassie Brilliant Blue R-250 and destained in 10% methanol and 7% acetic acid, until observation of proteolytic bands. Trypsin (2.5 mg/ml) was used as a positive control and parasite-free medium as negative control. For MS/MS analysis preparative gelatin-free gels stained with Coomassie Brilliant Blue were obtained following the same electrophoretic parameters in keratin-free conditions with 20 µl of the Bcl and Bcs samples.

2.6. Protein identification by LC-MS/MS

Bands from the preparative gels which matched with proteolytic bands on the analytic gels were manually excised for mass spectrometry identification. After excision, they were washed with 25 mM NH₄HCO₃–50% ethanol for 5 min at room temperature, then reduced 1 h at 56 °C with 10 mM DTT–50 mM NH₄HCO₃ and alkylated 30 min in darkness with 55 mM iodoacetamide–50 mM NH₄HCO₃. Finally, bands were washed with 25 mM NH₄HCO₃–5% acetonitrile (ACN) for 15 min and 25 mM NH₄HCO₃–50% ACN for 10 min. After dehydration in 100% ACN and SpeedVac drying, the volume band was evaluated and three volumes of a 10 ng/µl-trypsin (V511, Promega) solution in 25 mM NH₄HCO₃ were added. Digestion was performed at 37 °C for 5 h. After centrifugation, tryptic peptides were extracted by ACN, sonicated for 5 min and centrifuged. For mass spectrometry analyses, samples (6 µl) were first loaded and desalted on a trap (0.3-mm inner diameter × 5 mm) at 30 µl/min, 2% CH₃CN with 0.06% trifluoroacetic acid for 3 min and then separated on an analytical column (75-µm inner diameter × 15 cm) (both PepMap C₁₈, Dionex, Voisins Le Bretonneux, France) over a 45-min linear gradient of 10–90% CH₃CN, with 0.1% formic acid. The flow rate through the column was 200 nl/min. For ion trap MS, a LCQ Deca mass spectrometer with a nano electrospray interface (ThermoFinnigan, Les Ulis, France) was used. Ionization (1.8 kV ionization potential) was performed with a liquid junction and a non-coated capillary probe (New Objective, Cambridge, USA). Peptide ions were analyzed by the data-dependent “triple play” method: (i) full MS scan (*m/z* 400–1400), (ii) zoomscan (scan of the major ion with bigger resolution), and (iii) MS/MS of this ion. X!Tandem (downloaded from <http://www>.

the thegpm.org) was used for analyzing raw data, converted in mgf-file. The database, used for matching mass spectra with peptide sequences, was a local *Blastocystis* ST7 protein database deduced from genome data [13]. Parameters used in X!Tandem searches were the following: enzyme, trypsin; variable modifications, cysteine carbamidomethylation and methionine oxidation; maximum missed cleavages, 1; peptide mass tolerance ranging from -1.5 to $+1.5$ Da and 0.8 Da for fragment.

2.7. Western blot analysis

For immunological detection, proteins from Bcs were resolved in a 10%-polyacrylamide gel without reducing agents as in Section 2.5. They were then transferred onto polyvinylidene difluoride membranes (PVDF, Millipore) using a semi-dry system (Hoefer Semiphor, San Francisco, USA). Membranes were saturated in 5% skimmed milk in PBS and incubated overnight at 4°C with the anti-legumain monoclonal antibody Mab1D5 [16,17] at $1.2\ \mu\text{g}/\text{ml}$. After washing in 0.1% Triton X-100 in PBS, membranes were incubated for 1 h with HRP-conjugated anti-mouse immunoglobulin A, M, G diluted at 1:5000 (Invitrogen). Membranes were then revealed with ECL kit (ECL Plus Western blotting detection system, GE Healthcare).

3. Results

3.1. Prediction of *Blastocystis* ST7 secreted proteases

In order to identify virulence factor candidates, the whole genome sequence of *Blastocystis* ST7 [13] was searched for putative secreted proteases. During genome annotation, 18 were identified as probably secreted using SignalP3.0 and PSORT II algorithms [13]. In this study, we performed a second screening of putative secreted proteases using the same algorithms, and we enriched the repertoire with 7 further candidates (CBK22721.2, CBK21528.2, CBK22109.2, CBK22775.2, CBK21674.2, CBK20556.2, CBK23026.2) (Supplementary Table S1). Three out of the first 18 predicted ORFs/genes were not retained for further analysis since they either had no gene ontology (CBK22680.2 and CBK24206.2 ORFs) or possessed an inappropriate annotation (CBK22386.2 gene with a cellulase annotation). We also corrected the start methionine for the CBK20860.2, CBK21032.2 and CBK25399.2 genes, and introns for CBK21032.2 and CBK25399.2 according to protein size of their orthologs. Finally, our annotation strategy set a list of 22 putative secreted proteases which correspond to 20 cysteine proteases, 1 serine protease and 1 aspartic protease (Supplementary Table S1).

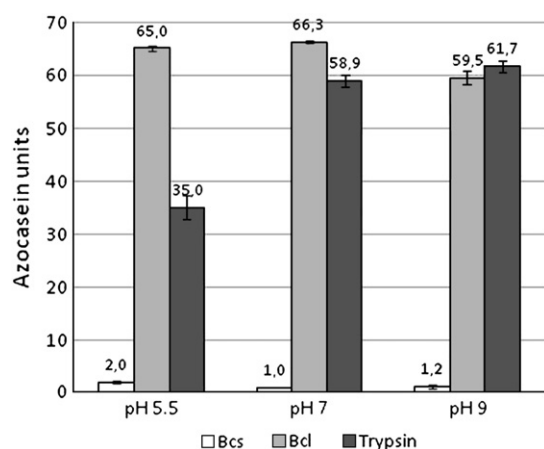


Fig. 1. Comparative analysis of proteolytic activities of *Blastocystis* ST7 culture supernatant (Bcs) and cell lysate (Bcl) at different pH. Error bars represent confidence interval at 95% which were calculated from a triplicate assay. Data correspond to one out of two independent biological repeats. A positive control for proteolytic activities was done using trypsin ($2.5\ \text{mg}/\text{ml}$).

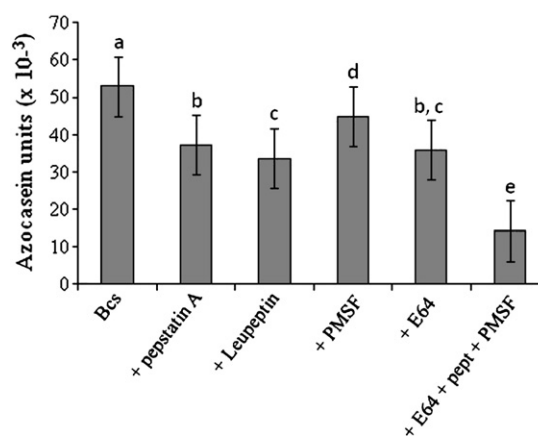


Fig. 2. Inhibition pattern of the *Blastocystis* ST7 culture supernatant (Bcs) protease activity. Error bars represent confidence interval at 95% which were calculated from a triplicate assay. Data correspond to one out of four independent biological repeats. Letters (a to e) refer to Student's tests, conditions labeled with different letters being significantly different. Protease inhibitors were pepstatin A ($100\ \mu\text{g}/\text{ml}$; aspartic protease inhibitor), leupeptin ($100\ \mu\text{g}/\text{ml}$; serine and cysteine protease inhibitor), PMSF ($1\ \text{mM}$; serine protease inhibitor) and E64 ($1\ \mu\text{M}$; cysteine protease inhibitor).

3.2. Presence of proteolytic activities in *Blastocystis* ST7 culture supernatant

In order to confirm the *in silico* prediction of secreted proteases, we evaluated and compared proteolytic activities of Bcs and Bcl extracts by azocasein assays. While Bcl sample exhibited an important protease activity with no pH dependency (probably due to a large diversity of intracellular proteases), the strongest activity for the Bcs one was observed at pH 5.5 (Fig. 1). Various protease inhibitors were then used at pH 5.5 to assess the nature of Bcs proteases (Fig. 2). The strongest proteolytic activity inhibition was noticed with leupeptin (36.5%), a serine and cysteine protease inhibitor. Strong inhibitions were also observed with the specific cysteine protease inhibitor E64 (32.1%) and the aspartic

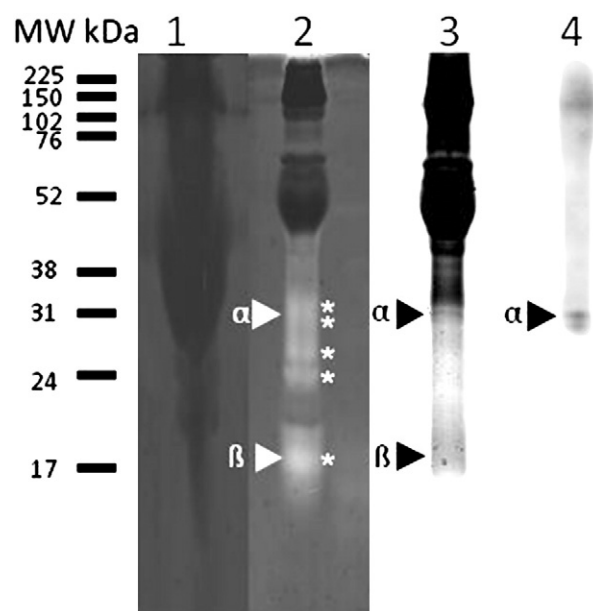


Fig. 3. Electrophoretic characterization of proteases in *Blastocystis* ST7 culture supernatant (Bcs). (1–2) The gelatin zymogram displayed no proteolytic band in a sterile medium sample (1) whereas five major ones (asterisks) were observed in Bcs (2). (3) Proteins from the Bcs extract were separated on a Coomassie Blue-stained SDS-PAGE gel for mass spectrometry identification of the proteolytic band contents. Bands α and β were excised for MS/MS analysis. (4) The anti-*Blastocystis* legumain monoclonal antibody Mab1D5 highlighted 2 bands of the Bcs extract by Western blot.

Table 1
MS/MS identification of the *Blastocystis* ST7 proteolytic α and β bands from parasite culture supernatant.

Band	Identified proteins				Experimental data		Matched peptides			
	Accession number	Function	Coverage %	Score (log(e))	MM (Da)	Δ MM (Da)	Sequence and position within the protein ^a	MM (Da)	Score (log(e))	e-Value
α	CBK21815.2	Legumain	8	12.1	2027.927 1555.744	–1.478 –0.992	⁸⁹ GVVVDYEGEDVTPENFMK ¹⁰⁶ ³⁴ HQADVAHAYQIMR ⁴⁶	2028.2 1539.7	–5 –2.9	6.73e-3 5.5e-2
β	CBK25506.2	Cathepsin B	5	5.8	1636.801	0.269	²⁷⁰ LVGWGVEDGVEYWK ²⁸³	1636.8	–5.9	2.73e-3

^a Underlined and in bold: oxidized methionine.

protease inhibitor pepstatin A (29.6%). A lower inhibition was observed using the serine protease inhibitor PMSF (15.1%). We also tested E64, pepstatin A and PMSF in combination, and observed an inhibition of 73%, suggesting the occurrence of other classes of proteases, or inhibitor-resistant proteases. Nevertheless we established that cysteine proteases are among the most active proteolytic enzymes characterized in Bcs in accordance with the *in silico* prediction. Each of the four repetitions of this experiment led to a similar inhibition pattern.

3.3. Identification of two cysteine proteases in *Blastocystis* ST7 culture supernatant

A gelatin zymographic SDS-PAGE was realized to identify some of the proteases characterized in the previous step (Fig. 3 lanes 1 and 2), and five major proteolytic bands were resolved. Only two of them (bands α and β of about 31 and 18 kDa, respectively) were Coomassie Blue-stained on SDS-PAGE gels (Fig. 3 lane 3), and were excised for MS/MS analysis.

Two peptides (Supplementary Fig. S1) from the α band (31 kDa) significantly matched with the amino acid sequence of the CBK21815.2 protein (398 aa, 45.2 kDa) which was annotated as the legumain we predicted to be secreted (Tables 1 and S1). This protein is identical to the *Blastocystis* ST7 legumain ACY95293.1 which was previously reported to be located at the parasite surface by Wu et al. [17]. The anti-*Blastocystis* legumain monoclonal antibody (MAb1D5) [16] highlighted two well-defined stained bands (Fig. 3 lane 4), at a molecular weight of approximately 31 kDa, which is consistent with the α band excised for mass spectrometry analysis (Fig. 3 lane 3) and with the work done by Wu et al. [17]. The size discrepancy between the experimental and theoretical molecular weights can be explained by a potential auto-cleavage at a carboxy-terminal asparagine residue as suggested by

Wu et al. [17], 7 Asn of the C-terminal region being candidates for such a cleavage site (Fig. 4A). This legumain was previously located at the parasite surface [17], probably within the cell coat that could constitute a transition step towards the secreted location as observed in our study.

One peptide (Supplementary Fig. S1) issued from the β band (18 kDa) corresponded to the amino acid sequence of cathepsin B (CBK25506.2, 321 aa, 35.7 kDa) (Tables 1 and S1). Since maturation (cleavage of the signal peptide and the propeptide (Fig. 4B)) does not seem to be sufficient to explain this size discrepancy, the truncated form we observed may rather result from autolysis or degradation by other proteases. A conserved 20-amino acid sequence corresponding to the cathepsin B occluding loop is predicted (Fig. 5), and should allow both endopeptidase and exopeptidase activities [18]. The asparaginyl residues present at positions 28, 37 and 62 within the *Blastocystis* ST7 cathepsin B could mark the end of the propeptide domain and represent targets for other proteases such as asparaginyl endopeptidases or legumains which are known to mature cathepsins [19,20]. We hypothesize that the secreted legumain may fill this role.

4. Discussion

Several enteric pathogens can modulate the protease balance which regulates the intestinal epithelial cell microenvironment inducing intestinal pathobiology. They can secrete specific proteases that stimulate gut secretion or inflammation, inhibit host immunity and, therefore, promote disease. As an example, the differences observed in term of virulence between *E. dispar* and *E. histolytica* is due to the secretion of specific proteases and to an increased protease activity for *E. histolytica* [5]. Interestingly, *Blastocystis* ST4 and ST7 proteolytic enzymes are able to increase the intestinal permeability,

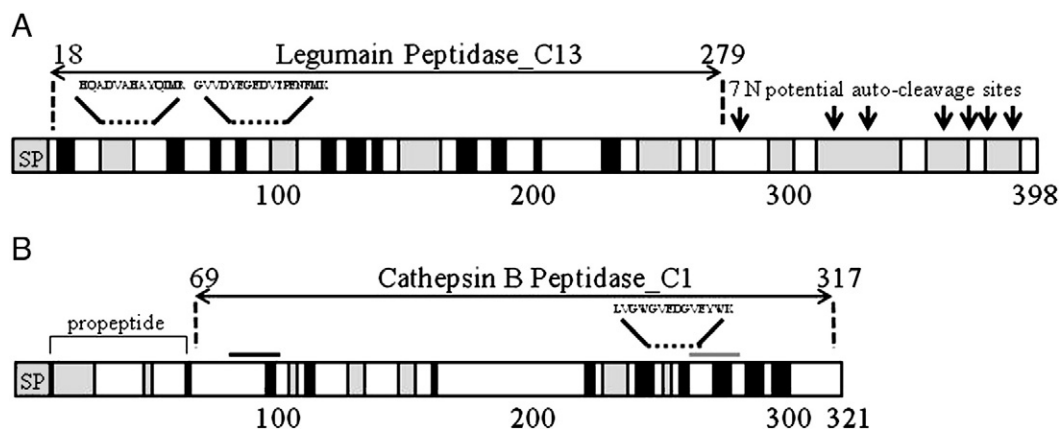


Fig. 4. Schematic representation of *Blastocystis* ST7 legumain (CBK21815.2, 398 aa) (A) and cathepsin B (CBK25506.2, 321 aa) (B) sequences. Full length proteins are represented with different predicted elements: signal peptide (SP), secondary structure consensus with beta strand in black and alpha helix in gray, and conserved Pfam domains (doubled arrow). The dotted lines correspond to the peptides identified by MS/MS. The legumain (A) presents 7C-terminal Asn residues (arrows) which constitute potential auto-cleavage sites. The cathepsin B (B) contains two active domains, a cysteine active site (black line) and a histidine one (gray line), as revealed by the MotifScan algorithm.

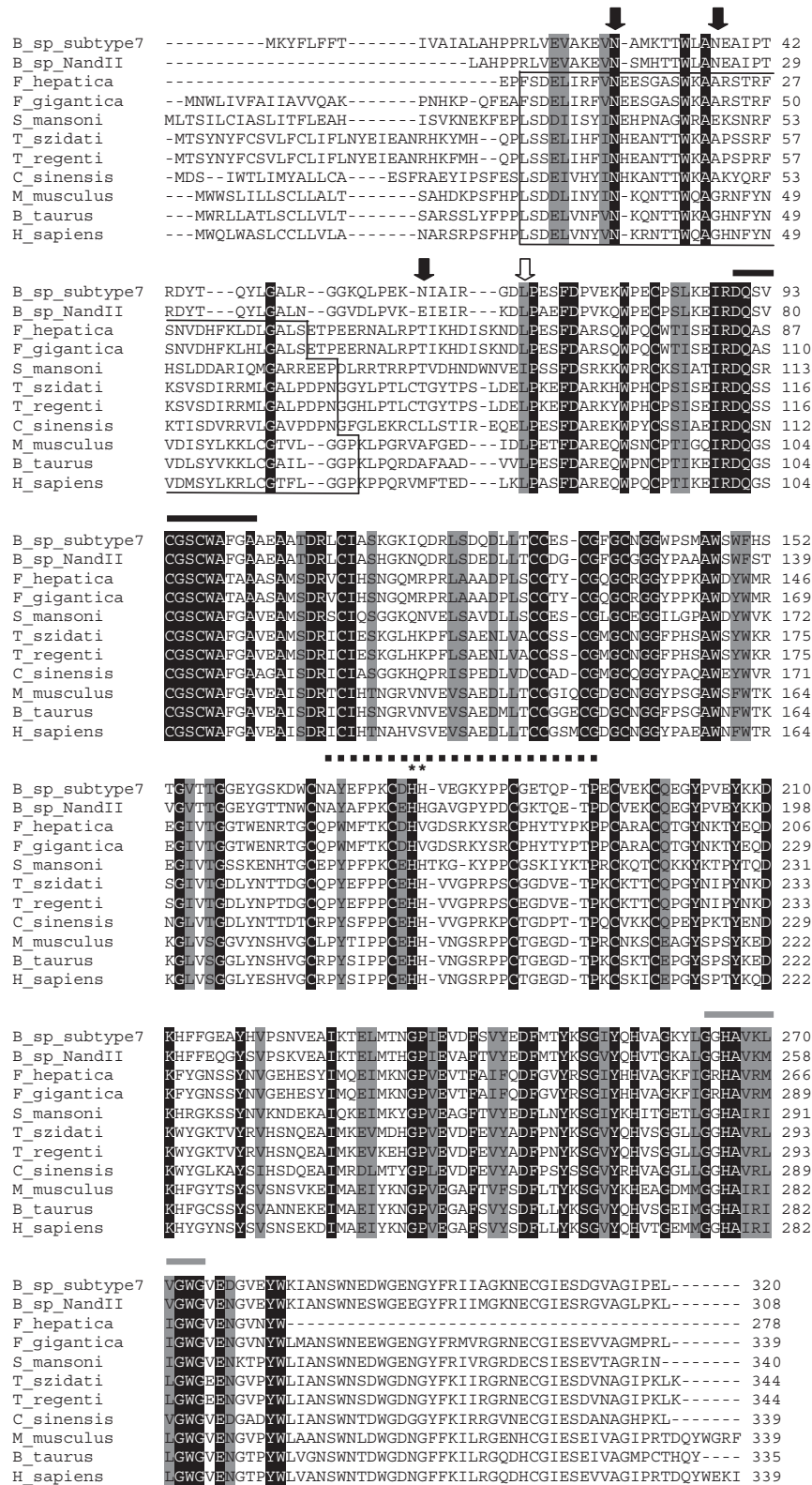


Fig. 5. Multiple sequence alignment of cathepsin B-like enzymes. ClustalW alignment includes the following: (i) the *Blastocystis* subtype 7 cathepsin B (CBK25506), and its counterpart from the *Blastocystis* NandII strain (ST1, BHL00000346, from the TBestDB database at (<http://amoebidia.bcm.umontreal.ca/pepdb/searches/login.php>)) to evaluate intra subtype variation; (ii) enzymes from the parasites *Fasciola hepatica* (ABU62925.1) and *Fasciola gigantica* (AA073002) since *Fasciola* is well studied in term of protease maturation (26); (iii) enzymes from the parasites *Schistosoma mansoni* (XP_002581167), *Trichobilharzia szidati* (ACG50797), *Trichobilharzia regenti* (ABS57370), and *Clonorchis sinensis* (ABM47070) which display higher identity with *Blastocystis* ST7 cathepsin according to MEROPS BLAST; (iv) enzymes from the mammals *Mus musculus* (AAA37375), *Bos taurus* (AAA03064) and *Homo sapiens* (AAC37547) which were used as an external group. Conserved residues and similar ones are shaded in black and gray, respectively. The conserved active sites are highlighted by a black line (cysteine active site) and by a gray line (histidine active site). The occluding loop is delineated by the dotted line and its two essential histidine residues labeled with asterisks. The white arrow marks the beginning of the peptidase C1 domain while the box surrounds the propeptide domain. The black arrows indicate the asparaginyl residues of *Blastocystis* ST7 which are potentially involved in propeptide maturation by legumain.

degrade human secretory immunoglobulins A, induce IL-8 production which causes inflammation, tissue damage and gastrointestinal disturbances [10,11]. A recent survey demonstrated that isolates from symptomatic and asymptomatic patients discard differential protease activities with specific proteases in the symptomatic isolates [21]. However, proteases involved in these processes have never been identified. This work provides two protease candidates for the observed effects induced by *Blastocystis* ST7. This is of great interest since *Blastocystis* spp. are assumed to be involved in intestinal inflammatory diseases [4,21,22]. When compared to asymptomatic controls, biopsies from IBS patients show an increased proteolytic activity, which involves serine proteases which can activate the protease-activated receptors-2 (PAR-2) [23]. PAR-2 are G-protein-coupled receptors highly expressed in the gastrointestinal tract, and activated not only by trypsinase from mast cells but also by luminal proteases such as trypsin and some pathogen proteases [23]. A serine protease is predicted in the secretome of *Blastocystis* ST7 and may target PAR-2 receptors of the host. However the epithelial cell surface is protected by the intestinal mucus that acts as a physical barrier [24]. The present work and the *Blastocystis* ST7 predicted secretome [13] show that proteases and glycosidases may be at the parasite–host interface and be involved in mucus degradation. The impairment of mucus may be the initial step to inflammatory and allergic disturbance caused by chronic exposure to luminal antigens. Pathogen cysteine proteases are known to degrade mucins of the intestinal epithelium providing nutrient resources and enabling an access to epithelial cells [7]. They have also the capacity to target enteric neurons by activating PAR receptors after epithelium disruption and paracellular permeability increase [25]. The two cysteine proteases we identified in the *Blastocystis* ST7 secretome may therefore be involved in these pathological aspects, particularly in IBS *Blastocystis*-positive patients.

To conclude, this study is the first identification of proteases secreted by the intestinal parasite *Blastocystis* ST7. Since protease balance of the intestinal tract is crucial to maintain a functional tissue, hijacking of this physiological process by enteric pathogens can promote intestinal disorders. Understanding the underlying mechanisms of this disturbance is primordial to decipher intestinal pathobiology. Moreover, cysteine proteases are well known to be immunogenic [18]. They could also be used to develop both diagnostic and virulence markers and constitute promising targets for chemotherapy [18].

Supplementary materials related to this article can be found online at doi:10.1016/j.parint.2012.02.007.

Acknowledgements

This work was funded by grants from “Ministère de l'Education, de la Recherche et de la Technologie” and the French National Center for Scientific Research (CNRS). We thank Christophe Chambon for mass spectrometry analyses which were performed at UR Microbiologie-QuaSA/Plateforme protéomique (INRA Theix, France), and David G. Biron for critical reading and English corrections of the manuscript.

References

- [1] Tan KS. New insights on classification, identification, and clinical relevance of *Blastocystis* spp. *Clinical Microbiology Reviews* 2008;21:639–65.
- [2] Tan KS, Mirza H, Teo JD, Wu B, Macary PA. Current views on the clinical relevance of *Blastocystis* spp. *Current Infectious Disease Reports* 2010;12:28–35.
- [3] Poirier P, Wawrzyniak I, Albert A, El Alaoui H, Delbac F, Livrelli V. Development and evaluation of a real-time PCR assay for detection and quantification of *Blastocystis* in human stool samples: a prospective study in patients with hematological malignancies. *Journal of Clinical Microbiology* 2011;49:975–83.
- [4] Yakoob J, Jafri W, Beg MA, Abbas Z, Naz S, Islam M, et al. *Blastocystis hominis* and *Dientamoeba fragilis* in patients fulfilling irritable bowel syndrome criteria. *Parasitology Research* 2010;107:679–84.
- [5] Que X, Reed SL. Cysteine proteinases and the pathogenesis of amebiasis. *Clinical Microbiology Reviews* 2000;13:196–206.
- [6] Buret AG. Mechanisms of epithelial dysfunction in giardiasis. *Gut* 2007;56:316–7.
- [7] Moncada D, Keller K, Chadee K. *Entamoeba histolytica* cysteine proteinases disrupt the polymeric structure of colonic mucin and alter its protective function. *Infection and Immunity* 2003;71:838–44.
- [8] Mirza H, Tan KS. *Blastocystis* exhibits inter- and intra-subtype variation in cysteine protease activity. *Parasitology Research* 2009;104:355–61.
- [9] Sio SW, Puthia MK, Lee AS, Lu J, Tan KS. Protease activity of *Blastocystis hominis*. *Parasitology Research* 2006;99:126–30.
- [10] Puthia MK, Vaithilingam A, Lu J, Tan KS. Degradation of human secretory immunoglobulin A by *Blastocystis*. *Parasitology Research* 2005;97:386–9.
- [11] Puthia MK, Lu J, Tan KS. *Blastocystis ratti* contains cysteine proteases that mediate interleukin-8 response from human intestinal epithelial cells in an NF-kappaB-dependent manner. *Eukaryotic Cell* 2008;7:435–43.
- [12] Puthia MK, Sio SW, Lu J, Tan KS. *Blastocystis ratti* induces contact-independent apoptosis, F-actin rearrangement, and barrier function disruption in IEC-6 cells. *Infection and Immunity* 2006;74:4114–23.
- [13] Denoed F, Roussel M, Noel B, Wawrzyniak I, Da Silva C, Diogon M, et al. Genome sequence of the stramenopile *Blastocystis*, a human anaerobic parasite. *Genome Biology* 2011;12:R29.
- [14] Ng GC, Tan KS. Colony growth as a step towards axenization of *Blastocystis* isolates. *Parasitology Research* 1999;85:678–9.
- [15] Ho LC, Singh M, Suresh G, Ng GC, Yap EH. Axenic culture of *Blastocystis hominis* in Iscove's modified Dulbecco's medium. *Parasitology Research* 1993;79:614–6.
- [16] Tan SW, Ho LC, Moe KT, Chen XQ, Ng GC, Yap EH, et al. Production and characterization of murine monoclonal antibodies to *Blastocystis hominis*. *International Journal of Parasitology* 1996;26:375–81.
- [17] Wu B, Yin J, Texier C, Roussel M, Tan KS. *Blastocystis* legumain is localized on the cell surface, and specific inhibition of its activity implicates a pro-survival role for the enzyme. *Journal of Biological Chemistry* 2010;285:1790–8.
- [18] Sajid M, McKerrrow JH. Cysteine proteases of parasitic organisms. *Molecular and Biochemical Parasitology* 2002;120:1–21.
- [19] Beckham SA, Law RH, Smooker PM, Quinsey NS, Caffrey CR, McKerrrow JH, et al. Production and processing of a recombinant *Fasciola hepatica* cathepsin B-like enzyme (FhcatB1) reveals potential processing mechanisms in the parasite. *Biological Chemistry* 2006;387:1053–61.
- [20] Sajid M, McKerrrow JH, Hansell E, Mathieu MA, Lucas KD, Hsieh I, et al. Functional expression and characterization of *Schistosoma mansoni* cathepsin B and its trans-activation by an endogenous asparaginyl endopeptidase. *Molecular and Biochemical Parasitology* 2003;131:65–75.
- [21] Abdel-Hameed DM, Hassanin OM. Protease activity of *Blastocystis hominis* subtype3 in symptomatic and asymptomatic patients. *Parasitology Research* 2011;109:321–7.
- [22] Stark D, van Hal S, Marriott D, Ellis J, Harkness J. Irritable bowel syndrome: a review on the role of intestinal protozoa and the importance of their detection and diagnosis. *International Journal of Parasitology* 2007;37:11–20.
- [23] Cenac N, Andrews CN, Holzhausen M, Chapman K, Cottrell G, Andrade-Gordon P, et al. Role for protease activity in visceral pain in irritable bowel syndrome. *The Journal of Clinical Investigation* 2007;117:636–47.
- [24] Lievin-Le Moal V, Servin AL. The front line of enteric host defense against unwelcome intrusion of harmful microorganisms: mucins, antimicrobial peptides, and microbiota. *Clinical Microbiology Reviews* 2006;19:315–37.
- [25] Steck N, Mueller K, Schemann M, Haller D. Bacterial proteases in IBD and IBS. *Gut* 2011, doi:10.1136/gutjnl-2011-300775.

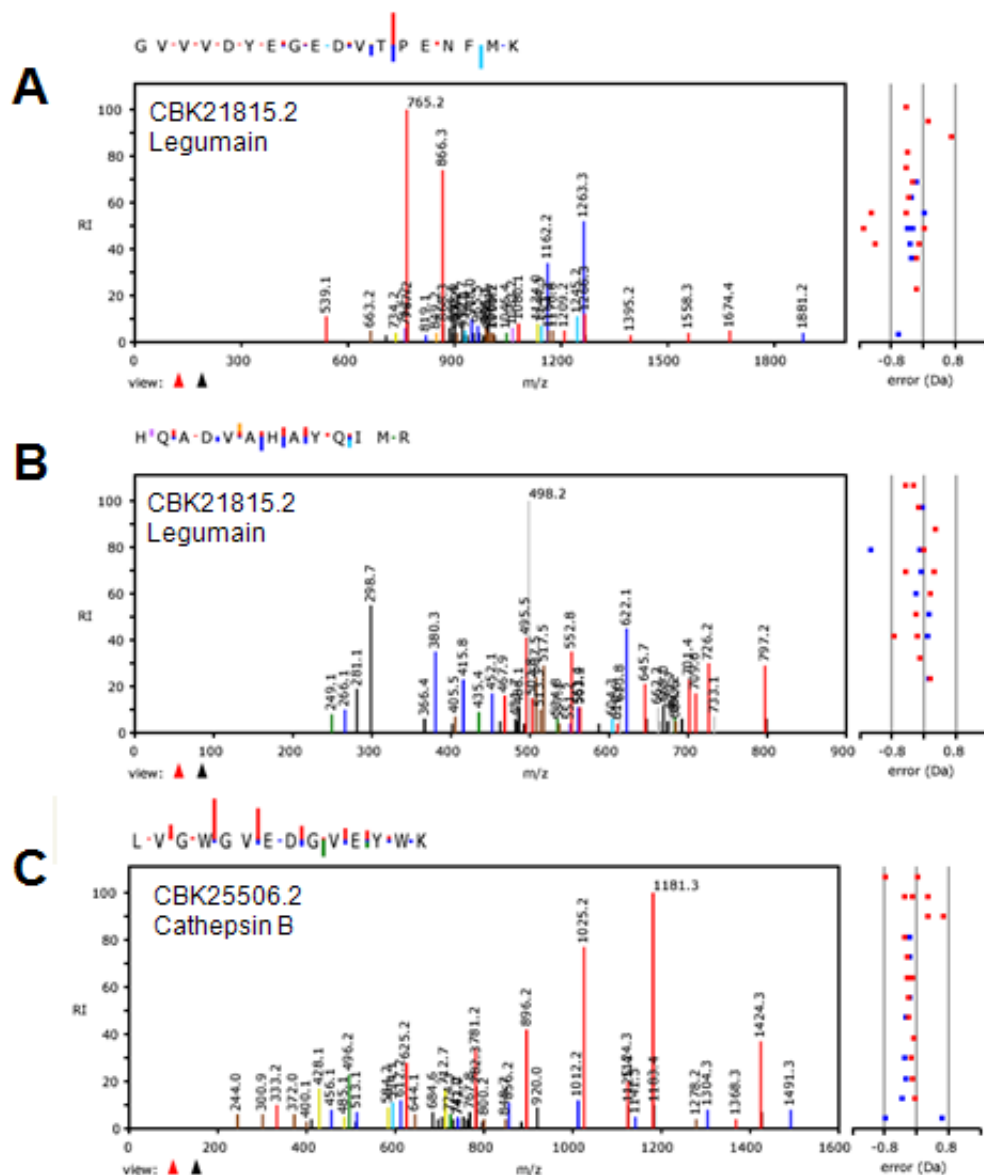


Figure S1 : MS/MS fragmentation spectra of tryptic peptides from the Bcs proteolytic bands α (A, B) and β (C). Tandem mass spectra were analyzed by X!tandem using the *Blastocystis* genome database. Mass indications label the +1 b and y ions from non modified fragments. Mass deviation between the experimental +1 b and y ions and the predicted ones is shown alongside the MS/MS fragmentation spectra. The deduced peptide sequence and protein identification (accession number and annotation) are shown on each spectrum.

Table S1: *In silico* identification of 22 putative secreted proteases from *Blastoystis* sp. subtype 7

Genbank number ¹	SignalP 3.0 ² ₃	Psport II ⁴	EC number ⁵	Enzyme annotation ⁵	MEROPS Family ⁶	Pfam domain ⁷	Gene Ontology ⁸	GO number ₈	EST ⁹	pI/MW (Da) ¹⁰
CBK22721.2	0.999	66.7 %: extracellular	EC 3.4.22.34	Legumain	C13	Peptidase C13	cysteine-type endopeptidase	GO:0004197	presence	5.23 / 48174
CBK25024.2	0.999	66.7 %: extracellular	EC 3.4.22.34	Legumain	C13	Peptidase C13	cysteine-type endopeptidase	GO:0004197	absence	7.64 / 41962
CBK20860.2	1.000	55.6 %: extracellular	EC 3.4.22.34	Legumain	C13	Peptidase C13	cysteine-type endopeptidase	GO:0004197	presence	5.15 / 50012
CBK21528.2	0.997	55.6 %: extracellular	EC 3.4.22.34	Legumain	C13	Peptidase C13	cysteine-type endopeptidase	GO:0004197	presence	4.99 / 48856
CBK21815.2	1.000	55.6 %: extracellular	EC 3.4.22.34	Legumain	C13	Peptidase C13	cysteine-type endopeptidase	GO:0004197	presence	4.97 / 45196
CBK22109.2	1.000	22.2 %: extracellular	EC 3.4.22.34	Legumain	C13	Peptidase C13	cysteine-type endopeptidase	GO:0004197	presence	5.07 / 48147
CBK22661.2	0.999	66.7 %: extracellular	EC 3.4.22.34	Legumain	C13	Peptidase C13	cysteine-type endopeptidase	GO:0004197	presence	5.23 / 47120
CBK22774.2	0.999	55.6 %: extracellular	EC 3.4.22.34	Legumain	C13	Peptidase C13	cysteine-type endopeptidase	GO:0004197	presence	4.78 / 47799
CBK22775.2	1.000	66.7 %: extracellular	EC 3.4.22.34	Legumain	C13	Peptidase C13	cysteine-type endopeptidase	GO:0004197	presence	4.78 / 47736
CBK20763.2	0.974	77.8 %: extracellular	EC 3.4.22.15	Cathepsin L	C1A	Inhibitor I29, Peptidase C1	cysteine-type endopeptidase	GO:0004197	presence	5.27 / 34259
CBK24846.2	0.995	66.7 %: extracellular	EC 3.4.22.15	Cathepsin L	C1A	Inhibitor I29, Peptidase C1	cysteine-type endopeptidase	GO:0004197	presence	5.32 / 34046
CBK21032.2	0.996	66.7 %: extracellular	EC 3.4.22.15	Cathepsin L	C1A	Inhibitor I29, Peptidase C1	cysteine-type endopeptidase	GO:0004197	presence	5.32 / 58940
CBK21674.2	0.996	77.8 %: extracellular	EC 3.4.14.1	Cathepsin C	C1A	Peptidase C1	cysteine-type endopeptidase	GO:0004197	presence	5.51 / 63771
CBK22330.2	1.000	66.7 %: extracellular	EC 3.4.14.1	Cathepsin C	C1A	Peptidase C1	cysteine-type endopeptidase	GO:0004197	presence	5.65 / 63733
CBK25399.2	0.823	66.7 %: extracellular	EC 3.4.14.1	Cathepsin C	C1A	Peptidase C1	cysteine-type endopeptidase	GO:0004197	presence	5.43 / 64988
CBK22745.2	0.999	77.8 %: extracellular	EC 3.4.22.1	Cathepsin B	C1A	Peptidase C1	cysteine-type peptidase	GO:0008234	absence	5.58 / 36352
CBK25506.2	0.998	55.6 %: extracellular	EC 3.4.22.1	Cathepsin B	C1A	Peptidase C1	cysteine-type endopeptidase	GO:0004197	presence	4.97 / 35652
CBK25507.2	0.998	44.4 %: extracellular	EC 3.4.22.1	Cathepsin B	C1A	Peptidase C1	cysteine-type endopeptidase	GO:0004197	presence	4.90 / 35675
CBK19711.2	0.999	55.6 %: extracellular	No match	No match	C1A	Peptidase C1	cysteine-type peptidase	GO:0008234	absence	6.09 / 71554
CBK20556.2	0.993	55.6 %: extracellular	EC 3.4.22.30	caricain	C1A	Inhibitor I29, Peptidase C1	cysteine-type endopeptidase	GO:0004197	presence	5.28 / 35850
CBK23053.2	1.000	44.4 %: extracellular	No match	No match	S8A	Peptidase S8	serine-type endopeptidase	GO:0004252	absence	6.17 / 68295
CBK23026.2	0.999	66.7 %: extracellular	No match	No match	A01	Aspartyl protease	aspartic-type endopeptidase	GO:0004190	absence	5.95 / 46416

¹ In red, putative secreted proteases from this work. In black, blue and green, putative secreted proteases from [23] which were either not refined (black), or corrected (this work) for the start methionine (blue) or for the start methionine and introns (green).

² The score is a probability that the protein possess a signal peptide

³ <http://www.cbs.dtu.dk/services/SignalP/>

⁴ <http://psort.hgc.jp/form2.html>

⁵ <http://www.expasy.org/>

⁶ <http://merops.sanger.ac.uk/>

⁷ <http://pfam.sanger.ac.uk/>

⁸ <http://www.geneontology.org/>

⁹ EST presence or absence resulting from EST data banks used for *Blastocystis* subtype 7 genome annotation

¹⁰ http://expasy.org/tools/pi_tool.html

Chapitre III.

Conclusions et perspectives

Le séquençage du génome de *Blastocystis* sp. ST7 a permis de disposer de données génomiques permettant d'apporter des éclaircissements concernant les mécanismes potentiels d'évolution mais également sur les répertoires géniques impliqués dans l'adaptation du parasite aux conditions d'anaérobiose du tractus intestinal et dans la pathogénie.

Si le séquençage et l'annotation de ce génome nous ont permis de poser des hypothèses sur la fonction et le rôle des gènes, ainsi que d'entrevoir les capacités de ce parasite, ces résultats issus d'analyses *in silico* restent des prédictions obtenues par l'utilisation de logiciels bioinformatiques et sont donc soumis aux limites des systèmes actuels d'annotation automatique. Ainsi, 252 gènes de *Blastocystis* sp. ST7, soit 4% du génome, ne présentent aucune homologie avec les séquences présentes dans les bases de données et 1215 gènes, soit près de 20% du génome, possèdent une annotation discutable en raison d'un faible pourcentage d'identité avec leurs orthologues potentiels et d'une E-value supérieur à 1. Dans le cas de *Blastocystis* sp. ST7, nous avons réalisé une annotation par domaine, ce qui ne permet pas de déterminer avec précision la fonction des gènes, notamment lors de la présence de plusieurs domaines. Ainsi plusieurs protéines sont multifonctionnelles et participent donc à différentes fonctions biologiques. De même, certaines enzymes présentant des homologies de séquence ont des fonctions différentes. Ces raisons justifient la prudence d'une annotation par domaine et illustrent la difficulté d'attribuer une fonction précise sur la base d'une séquence (Schaechter et Ingraham, 2002). Dès lors, il est important de ne pas oublier que ces analyses restent des prédictions permettant d'orienter les expérimentations en laboratoire, et qu'elles ne remplaceront jamais les études expérimentales. Malgré tout, le séquençage d'un génome complet est une stratégie intéressante, notamment quand on ne dispose pas de beaucoup de données biologiques sur un organisme.

I. Un génome qui a évolué

Lors de l'analyse du génome de *Blastocystis* sp. ST7, nous avons suggéré l'existence de plusieurs mécanismes d'évolution. En effet, ce génome présente des blocs de segments paralogues témoignant d'un phénomène de duplication. S'il nous est impossible de trancher entre les deux hypothèses pouvant expliquer l'origine de ces duplications (duplication complète du génome ou duplications segmentaires), ce génome se caractérise par un taux élevé de réarrangements. Le projet de génomique comparative, initié en collaboration avec l'Institut Pasteur de Lille, la société Genoscreen et l'Academic Medical Center d'Amsterdam, qui a pour objectif de comparer quatre nouveaux génomes de *Blastocystis* spp. (STs 1, 4, 6 et 8), pourrait permettre de déterminer si ces duplications et le taux élevé de réarrangement sont caractéristiques du genre *Blastocystis* ou spécifique du ST7. L'identification de familles de gènes dupliqués chez les différents STs pourrait également permettre de mieux comprendre en quoi ces duplications ont pu concourir à l'adaptation de *Blastocystis* spp. à son environnement.

Le génome de *Blastocystis* sp. ST7 semble également avoir évolué avec l'apport de nouveaux gènes via des transferts horizontaux. En effet, un nombre conséquent de gènes ont probablement pour origine des bactéries ou des archaebactéries. Les analyses phylogénétiques ne sont pas assez résolutes pour permettre d'identifier les donneurs, suggérant ainsi des transferts horizontaux anciens ou une évolution rapide de ces gènes. Certains de ces gènes ont pu être également hérités d'eucaryotes évoluant dans l'environnement intestinal. Ces observations montrent l'importance de ce mécanisme en termes d'adaptation. Ce mécanisme de transfert horizontal, déjà mis en évidence chez le straménopile *Phaeodactylum tricornutum* (Bowler et al., 2008), a permis à cette diatomée d'acquérir des processus de réparation de l'ADN ou la capacité de synthèse d'une coque constituée de silice. Il est probable que ce processus ait permis à *Blastocystis* sp. ST7 d'acquérir des gènes (dehydrogénase, MFS (Major facilitator

transporter), cadhérine) lui permettant de s'adapter et de survivre aux conditions du tractus intestinal.

Dans l'objectif de mieux décrire et comprendre ces transferts horizontaux, le projet de génomique comparative pourrait permettre de déterminer un coeur de gènes, présent dans les cinq génomes. Il pourrait également permettre de confirmer l'acquisition de gènes issus de transfert horizontal dans la lignée de *Blastocystis* spp. Les gains de gènes peuvent en effet être corrélés à des mécanismes d'adaptations spécifiques. L'existence de différences en termes d'acquisition de gènes entre les différents sous-types de *Blastocystis* spp. pourraient aider à classer phylogénétiquement ces sous-types entre eux.

II. Un métabolisme adapté aux conditions d'anoxie du tractus intestinal

Il est vraisemblable que *Blastocystis* spp. possède des mitochondries anaérobies capables de faire une respiration fumarate à l'image de ce qui existe chez les helminthes (Tielens et al., 2002). En effet *Blastocystis* sp. ST7 possède les complexe I et II de la phosphorylation oxydative, mais est dépourvu des complexes III, IV et V. Les complexes I et II peuvent fonctionner avec certaines enzymes du cycle de Krebs et permettre ainsi le transfert des électrons au fumarate, produisant ainsi du succinate. Des résultats préliminaires, en RMN, ont montré des produits finaux caractéristiques de cette respiration fumarate (données personnelles). Il convient maintenant de reproduire ces analyses et de rechercher la nature des quinones (notamment la rholoquinone) dans cette chaîne respiratoire ce qui serait un indicateur de cette respiration fumarate. De plus la prédiction du protéome des MLOs a mis en évidence la présence d'enzymes hydrogénomales (la PFOR et l'hydrogénase à Fer), dont nous avons pu confirmer la localisation expérimentalement. Nous avons également mis en évidence les sous-unités α et β de la succinyl-CoA synthétase impliquées dans la production d'énergie à la fois chez les hydrogénosomes et les mitochondries. De même nous avons prédits plusieurs enzymes pouvant prendre en charge le pyruvate (une pyruvate déshydrogénase, une PFOR, une PNO), ce qui suggère la présence de plusieurs voies permettant de synthétiser de l'acétyl-CoA à partir du pyruvate. L'ensemble de ces données conduit à penser que les MLOs de *Blastocystis* sp. ST7 représentent une situation particulière parmi les différents types d'organites apparentés aux mitochondries, ayant permis une adaptation réussie du parasite au tractus intestinal des animaux dont l'homme. Il reste maintenant à décrire expérimentalement la composition de cet organite et son fonctionnement de façon plus précise.

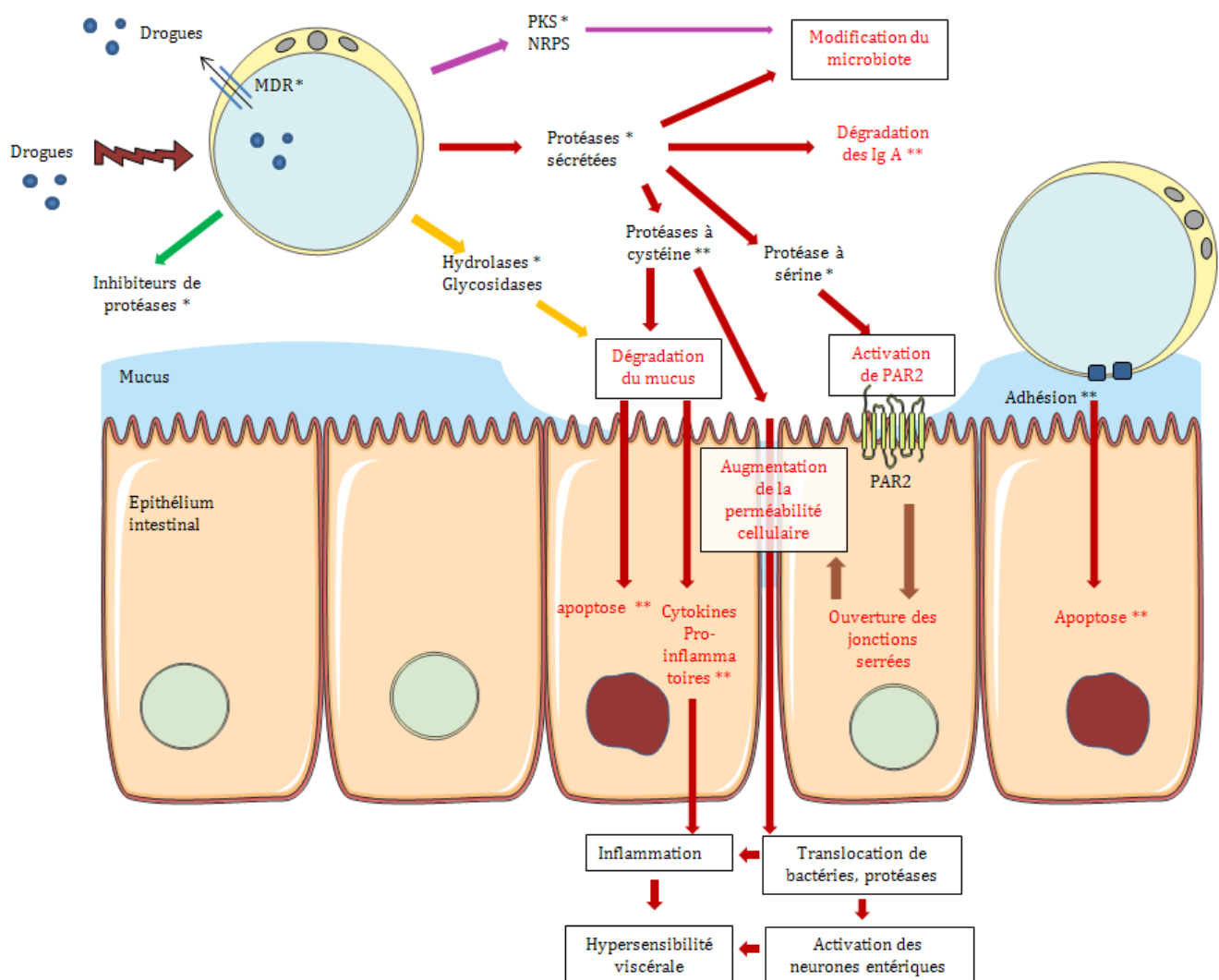


Figure 22 : Modèle de mécanismes de pathogénie chez *Blastocystis* sp. ST7. L'infection par *Blastocystis* spp. pourrait entraîner plusieurs dysfonctionnements intestinaux comme la dégradation des IgA (Puthia et al., 2005), l'entrée en apoptose des cellules épithéliales (Puthia et al., 2006), l'augmentation de la perméabilité cellulaire, ou l'induction de la production d'interleukines (Long et al, 2001, Puthia et al., 2005, 2008). La dégradation du mucus pourrait être engendrée par la sécrétion par le parasite de protéases, d'hydrolases ou de glycosidases. Les protéases seraient également à l'origine de la dégradation des IgA, de l'entrée en apoptose des cellules intestinales, de la rupture des jonctions cellulaires de l'épithélium et de l'induction de la production d'IL-8. Elles modifieraient probablement aussi le microbiote au même titre que les PKS et NRPS. La protéase à sérine pourrait être impliquée dans l'activation des récepteurs PAR2, entraînant l'ouverture des jonctions serrées. Les inhibiteurs de protéases sécrétés joueraient probablement un rôle de protection contre les protéases de l'hôte. Les MDR pourraient expliquer les phénomènes de résistance aux drogues. Quelques protéines potentiellement impliquées dans l'adhésion ont également été identifiées.

* : données génomiques

** : données expérimentales (ce travail et les références citées ci-dessus)

III. Un réel potentiel pathogène

Les prédictions *in silico* réalisées sur le génome de *Blastocystis* sp. ST7, nous ont permis d'identifier plusieurs gènes pouvant jouer un rôle dans la virulence du parasite. (Figure 22). Ainsi, certains gènes codent des protéines (cadhérines, protéines avec un domaine immunoglobuline) ayant des orthologues impliqués dans le processus d'adhésion. Il est tout à fait envisageable que ce soit un moyen de persistance du parasite dans l'organisme, à travers l'adhésion au mucus. Nous avons également mis en évidence des hydrolases qui pourrait attaquer les tissus de l'hôte afin de fournir les nutriments nécessaires à son développement. Les glycosidases, et les protéases prédites dans le sécrétome pourraient fournir une source de sucres et de protéines pour le parasite par dégradation du mucus. Il est envisageable que la détérioration de cette barrière mucoale facilite l'accès des protéases bactériennes ou parasitaires aux cellules épithéliales, entraînant une perturbation des jonctions serrées et de fait une augmentation de la perméabilité cellulaire de l'épithélium intestinal. De plus les protéases pourraient également induire l'entrée en apoptose des cellules épithéliales. En effet, l'entrée en apoptose et la rupture de jonctions serrées sont induites *in vitro* par contact avec *Blastocystis* sp. ST4 (Puthia et al., 2006). Les protéases de *Blastocystis* sp. ont également été suspectées d'induire des effets pathologiques sur des cellules intestinales humaines (Puthia et al., 2008; Tan et al., 2010), et sont capable d'induire la production d'IL-8 (Puthia et al., 2008). Leur implication dans la modulation ou la suppression des défenses immunitaires a aussi été prouvés expérimentalement chez *Blastocystis* sp. (Long et al., 2001; Puthia et al., 2005; Puthia et al., 2008; Tan et al., 2010).

Nous avons prédits dans le génome de *Blastocystis* sp. ST7 22 protéases sécrétées. Nous avons identifié deux protéases à cystéine (une cathepsine de type B et une légumine) présentes dans les surnageants de culture. Il serait intéressant de déterminer l'implication potentielle de ces deux protéases dans la physiopathologie de *Blastocystis* sp. ST7, et plus particulièrement de déterminer leur capacité à

compromettre l'intégrité des cellules intestinales et l'organisation de leur cytosquelette, y compris les protéines impliquées dans les jonctions serrées. A cette fin, le clonage et l'expression de ces protéases est en cours de réalisation. L'évaluation de leur implication dans la physiopathologie est prévue par mesure de la résistance trans-épithéliale et par immuno-marquage des protéines du cytosquelette : actine, myosine, et protéines des jonctions serrées (ZO-1, occludine). Deux lignées cellulaires intestinales humaines seront utilisées à ces fins : d'une part les cellules différenciées Caco2 et d'autre part les cellules I407 issues du jéjunum et de l'iléon. Il serait aussi intéressant d'identifier d'autres protéases pouvant être présentes dans les surnageants de culture. Une approche par analyses zymographiques sur gel en 2 dimensions devrait pouvoir permettre une meilleure résolution des protéases présentes dans les surnageants de culture. Les spots ainsi obtenus pourront être analysés par spectrométrie de masse.

Il a également été montré une prévalence plus importante du parasite chez des patients IBS (Giacometti et al., 1999). Des analyses de biopsies issues de patients IBS ont montré une augmentation de la perméabilité cellulaire associée à une perturbation des jonctions serrées (Piche et al., 2009; Bertiaux-Vandaele et al., 2011). Il a été proposé que les récepteurs de type 2 activés par les protéases (PAR-2) soient impliqués dans cette augmentation de la perméabilité (Bueno et Fioramonti, 2008). En effet, ces récepteurs PAR-2, activés par des protéases à sérine, jouent un rôle fondamental dans des pathologies digestives *via* notamment la redistribution des jonctions serrées. Il est envisagé de tester l'effet de la protéase à sérine prédite dans le sécrétome de *Blastocystis* sp. ST7 sur l'activation des récepteurs PAR-2 par PCR quantitative en suivant ces derniers. On peut aussi envisager que les protéases sécrétées puissent modifier l'environnement direct du parasite, notamment en perturbant et/ou dégradant la flore bactérienne pour se nourrir, pour se protéger ou simplement favoriser son développement. La présence de PKS et de NRPS pouvant synthétiser des toxines ou des antibiotiques appuie l'hypothèse que *Blastocystis* sp. ST7 peut modifier le microbiote qui

l'entoure. Cette dysbiose est un phénomène connu dans les cas d'IBS (Parkes et al., 2008).

Cet arsenal de gènes mis en évidence pourrait expliquer les cas de patients symptomatiques positifs à *Blastocystis* spp. La comparaison des différents génomes de *Blastocystis* sp. en cours de séquençage devrait pouvoir apporter des pistes sur le caractère plus ou moins pathogène de certains sous-types. En effet des différences dans certaines familles de gènes codant de potentiels facteurs de virulence, en termes de présence de gènes ou d'expansion de ces familles, et en particulier dans les répertoires dynamiques des gènes effecteurs dans les mécanismes de défense ou de virulence, pourraient être à l'origine de ces différences de pathogénie suggérées entre les sous-types de *Blastocystis* sp.

Au vu de ces différents résultats il apparaît plausible que *Blastocystis* sp. ST7 puisse être un agent pathogène. Les études en cours ou prévues devraient permettre d'avancer sur l'impact réel de ce parasite en santé humaine et dans le cas d'un rôle significatif de *Blastocystis* sp. d'alerter les responsables de santé publique sur la nécessité de réévaluer ce parasite.

Les protéases étant essentielles pour de nombreux parasites, elles constituent des cibles intéressantes pour le développement de nouvelles approches thérapeutiques (Smooker et al., 2010) ou vaccinales (Sajid et McKerrow, 2002). Des inhibiteurs de protéases ayant une efficacité démontrée contre des protéases parasitaires (North, 1994) seront évalués sur leur capacité à inhiber spécifiquement les protéases de *Blastocystis* spp. Ces enzymes pourront également servir en sérodiagnostic, des séra de patients symptomatiques et asymptomatiques seront analysés par Western Blot sur les protéases recombinantes pour évaluer leur signification clinique.

Pour conclure, mon travail de thèse montre la présence de facteurs de virulence potentiels dans le génome de *Blastocystis* sp. ST7 ainsi que l'expression de protéines qui, à l'interface hôte-parasite, pourraient altérer le fonctionnement intestinal de l'hôte. Ce travail fournit donc une base pour l'étude de la pathogénie de *Blastocystis* sp., et ouvre

de nombreuses pistes intéressantes à explorer pour évaluer l'impact de ce parasite en santé humaine.

Chapitre IV

Références bibliographiques

- Abe, N., Nagoshi, M., Takami, K., Sawano, Y., Yoshikawa, H., 2002. A survey of *Blastocystis* sp. in livestock, pets, and zoo animals in Japan. *Vet Parasitol* 106, 203-212.
- Abe, N., 2004. Molecular and phylogenetic analysis of *Blastocystis* isolates from various hosts. *Vet Parasitol* 120, 235-242.
- Abrahamsen, M.S., Templeton, T.J., Enomoto, S., Abrahante, J.E., Zhu, G., Lancto, C.A., Deng, M., Liu, C., Widmer, G., Tzipori, S., Buck, G.A., Xu, P., Bankier, A.T., Dear, P.H., Konfortov, B.A., Spriggs, H.F., Iyer, L., Anantharaman, V., Aravind, L., Kapur, V., 2004. Complete genome sequence of the apicomplexan, *Cryptosporidium parvum*. *Science* 304, 441-445.
- Alexander, J., Coombs, G.H., Mottram, J.C., 1998. *Leishmania mexicana* cysteine proteinase-deficient mutants have attenuated virulence for mice and potentiate a Th1 response. *J Immunol* 161, 6794-6801.
- Alvarez-Sanchez, M.E., Avila-Gonzalez, L., Becerril-Garcia, C., Fattel-Facenda, L.V., Ortega-Lopez, J., Arroyo, R., 2000. A novel cysteine proteinase (CP65) of *Trichomonas vaginalis* involved in cytotoxicity. *Microb Pathog* 28, 193-202.
- Anderson, I.J., Loftus, B.J., 2005. *Entamoeba histolytica*: observations on metabolism based on the genome sequence. *Exp Parasitol* 110, 173-177.
- Apic, G., Ignjatovic, T., Boyer, S., Russell, R.B., 2005. Illuminating drug discovery with biological pathways. *FEBS Lett* 579, 1872-1877.
- Arisue, N., Hashimoto, T., Yoshikawa, H., Nakamura, Y., Nakamura, G., Nakamura, F., Yano, T.A., Hasegawa, M., 2002. Phylogenetic position of *Blastocystis hominis* and of stramenopiles inferred from multiple molecular sequence data. *J Eukaryot Microbiol* 49, 42-53.
- Arisue, N., Hashimoto, T., Yoshikawa, H., 2003. Sequence heterogeneity of the small subunit ribosomal RNA genes among *blastocystis* isolates. *Parasitology* 126, 1-9.
- Armbrust, E.V., Berges, J.A., Bowler, C., Green, B.R., Martinez, D., Putnam, N.H., Zhou, S., Allen, A.E., Apt, K.E., Bechner, M., Brzezinski, M.A., Chaal, B.K., Chiovitti, A., Davis, A.K., Demarest, M.S., Detter, J.C., Glavina, T., Goodstein, D., Hadi, M.Z., Hellsten, U., Hildebrand, M., Jenkins, B.D., Jurka, J., Kapitonov, V.V., Kroger, N., Lau, W.W., Lane, T.W., Larimer, F.W., Lippmeier, J.C., Lucas, S., Medina, M., Montsant, A., Obornik, M., Parker, M.S., Palenik, B., Pazour, G.J., Richardson, P.M., Rynearson, T.A., Saito, M.A., Schwartz, D.C., Thamtrakoln, K., Valentin, K., Vardi, A., Wilkerson, F.P., Rokhsar, D.S., 2004. The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science* 306, 79-86.
- Arroyo, R., Alderete, J.F., 1995. Two *Trichomonas vaginalis* surface proteinases bind to host epithelial cells and are related to levels of cytoadherence and cytotoxicity. *Arch Med Res* 26, 279-285.
- Aslam, A., Quinn, P., McIntosh, R.S., Shi, J., Ghumra, A., McKerrow, J.H., Bunting, K.A., Dunne, D.W., Doenhoff, M.J., Morrison, S.L., Zhang, K., Pleass, R.J., 2008. Proteases from *Schistosoma mansoni* cercariae cleave IgE at solvent exposed interdomain regions. *Mol Immunol* 45, 567-574.
- Atkinson, H.J., Babbitt, P.C., Sajid, M., 2009. The global cysteine peptidase landscape in parasites. *Trends Parasitol* 25, 573-581.
- Aurrecoechea, C., Barreto, A., Brestelli, J., Brunk, B.P., Caler, E.V., Fischer, S., Gajria, B., Gao, X., Gingle, A., Grant, G., Harb, O.S., Heiges, M., Iodice, J., Kissinger, J.C., Kraemer, E.T., Li, W., Nayak, V., Pennington, C., Pinney, D.F., Pitts, B., Roos, D.S., Srinivasamoorthy, G., Stoeckert, C.J., Jr., Treatman, C., Wang, H., 2011. AmoebaDB and MicrosporidiaDB: functional genomic resources for Amoebozoa and Microsporidia species. *Nucleic Acids Res* 39, D612-619.
- Baggiolini, M., Loetscher, P., Moser, B., 1995. Interleukin-8 and the chemokine family. *Int J Immunopharmacol* 17, 103-108.
- Banaticla, J.E., Rivera, W.L., 2011. Detection and subtype identification of *Blastocystis* isolates from wastewater samples in the Philippines. *J Water Health* 9, 128-137.
- Barret, H., P., 1921. A method for the cultivation of *Blastocystis*. *Ann of trop med parasitol* 15, 113-116.
- Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D.C., Lennard, N.J., Caler, E., Hamlin, N.E., Haas, B., Bohme, U., Hannick, L., Aslett, M.A., Shallom, J., Marcello, L.,

- Hou, L., Wickstead, B., Alsmark, U.C., Arrowsmith, C., Atkin, R.J., Barron, A.J., Bringaud, F., Brooks, K., Carrington, M., Cherevach, I., Chillingworth, T.J., Churcher, C., Clark, L.N., Corton, C.H., Cronin, A., Davies, R.M., Doggett, J., Djikeng, A., Feldblyum, T., Field, M.C., Fraser, A., Goodhead, I., Hance, Z., Harper, D., Harris, B.R., Hauser, H., Hostetler, J., Ivens, A., Jagels, K., Johnson, D., Johnson, J., Jones, K., Kerhornou, A.X., Koo, H., Larke, N., Landfear, S., Larkin, C., Leech, V., Line, A., Lord, A., Macleod, A., Mooney, P.J., Moule, S., Martin, D.M., Morgan, G.W., Mungall, K., Norbertczak, H., Ormond, D., Pai, G., Peacock, C.S., Peterson, J., Quail, M.A., Rabinowitsch, E., Rajandream, M.A., Reitter, C., Salzberg, S.L., Sanders, M., Schobel, S., Sharp, S., Simmonds, M., Simpson, A.J., Tallon, L., Turner, C.M., Tait, A., Tivey, A.R., Van Aken, S., Walker, D., Wanless, D., Wang, S., White, B., White, O., Whitehead, S., Woodward, J., Wortman, J., Adams, M.D., Embley, T.M., Gull, K., Ullu, E., Barry, J.D., Fairlamb, A.H., Opperdoes, F., Barrell, B.G., Donelson, J.E., Hall, N., Fraser, C.M., Melville, S.E., El-Sayed, N.M., 2005. The genome of the African trypanosome *Trypanosoma brucei*. *Science* 309, 416-422.
- Bertiaux-Vandaele, N., Youmba, S.B., Belmonte, L., Lecleire, S., Antonietti, M., Gourcerol, G., Leroi, A.M., Dechelotte, P., Menard, J.F., Ducrotte, P., Coeffier, M., 2011. The expression and the cellular distribution of the tight junction proteins are altered in irritable bowel syndrome patients with differences according to the disease subtype. *Am J Gastroenterol* 106, 2165-2173.
- Bohm-Gloning, B., Knobloch, J., Walderich, B., 1997. Five subgroups of *Blastocystis hominis* from symptomatic and asymptomatic patients revealed by restriction site analysis of PCR-amplified 16S-like rDNA. *Trop Med Int Health* 2, 771-778.
- Bontempi, E., Cazzulo, J.J., 1990. Digestion of human immunoglobulin G by the major cysteine proteinase (cruzipain) from *Trypanosoma cruzi*. *FEMS Microbiol Lett* 58, 337-341.
- Boorom, K.F., Smith, H., Nimri, L., Viscogliosi, E., Spanakos, G., Parkar, U., Li, L.H., Zhou, X.N., Ok, U.Z., Leelayoova, S., Jones, M.S., 2008. Oh my aching gut: irritable bowel syndrome, *Blastocystis*, and asymptomatic infection. *Parasit Vectors* 1, 40.
- Boreham, P.F., Stenzel, D.J., 1993. *Blastocystis* in humans and animals: morphology, biology, and epizootiology. *Adv Parasitol* 32, 1-70.
- Bowler, C., Allen, A.E., Badger, J.H., Grimwood, J., Jabbari, K., Kuo, A., Maheswari, U., Martens, C., Maumus, F., Otilar, R.P., Rayko, E., Salamov, A., Vandepoele, K., Beszteri, B., Gruber, A., Heijde, M., Katinka, M., Mock, T., Valentin, K., Verret, F., Berges, J.A., Brownlee, C., Cadoret, J.P., Chiovitti, A., Choi, C.J., Coesel, S., De Martino, A., Detter, J.C., Durkin, C., Falcioratore, A., Fournet, J., Haruta, M., Huysman, M.J., Jenkins, B.D., Jiroutova, K., Jorgensen, R.E., Joubert, Y., Kaplan, A., Kroger, N., Kroth, P.G., La Roche, J., Lindquist, E., Lommer, M., Martin-Jezequel, V., Lopez, P.J., Lucas, S., Mangogna, M., McGinnis, K., Medlin, L.K., Montsant, A., Oudot-Le Secq, M.P., Napoli, C., Obornik, M., Parker, M.S., Petit, J.L., Porcel, B.M., Poulsen, N., Robison, M., Rychlewski, L., Ryneerson, T.A., Schmutz, J., Shapiro, H., Siaut, M., Stanley, M., Sussman, M.R., Taylor, A.R., Vardi, A., von Dassow, P., Vyverman, W., Willis, A., Wyrwicz, L.S., Rokhsar, D.S., Weissenbach, J., Armbrust, E.V., Green, B.R., Van de Peer, Y., Grigoriev, I.V., 2008. The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* 456, 239-244.
- Bruchhaus, I., Jacobs, T., Leippe, M., Tannich, E., 1996. *Entamoeba histolytica* and *Entamoeba dispar*: differences in numbers and expression of cysteine proteinase genes. *Mol Microbiol* 22, 255-263.
- Bruchhaus, I., Loftus, B.J., Hall, N., Tannich, E., 2003. The intestinal protozoan parasite *Entamoeba histolytica* contains 20 cysteine protease genes, of which only a small subset is expressed during in vitro cultivation. *Eukaryot Cell* 2, 501-509.
- Brumpt, E., 1912. *Blastocystis hominis* n. sp. et formes voisines. *Bull. Soc. Pathol. Exot* 5, 725-730.
- Bueno, L., Fioramonti, J., 2008. Protease-activated receptor 2 and gut permeability: a review. *Neurogastroenterol Motil* 20, 580-587.

- Carlton, J.M., Angiuoli, S.V., Suh, B.B., Kooij, T.W., Pertea, M., Silva, J.C., Ermolaeva, M.D., Allen, J.E., Selengut, J.D., Koo, H.L., Peterson, J.D., Pop, M., Kosack, D.S., Shumway, M.F., Bidwell, S.L., Shallom, S.J., van Aken, S.E., Riedmuller, S.B., Feldblyum, T.V., Cho, J.K., Quackenbush, J., Sedegah, M., Shoaibi, A., Cummings, L.M., Florens, L., Yates, J.R., Raine, J.D., Sinden, R.E., Harris, M.A., Cunningham, D.A., Preiser, P.R., Bergman, L.W., Vaidya, A.B., van Lin, L.H., Janse, C.J., Waters, A.P., Smith, H.O., White, O.R., Salzberg, S.L., Venter, J.C., Fraser, C.M., Hoffman, S.L., Gardner, M.J., Carucci, D.J., 2002. Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii yoelii*. *Nature* 419, 512-519.
- Carlton, J.M., 2003. Genome sequencing and comparative genomics of tropical disease pathogens. *Cell Microbiol* 5, 861-873.
- Carlton, J.M., Hirt, R.P., Silva, J.C., Delcher, A.L., Schatz, M., Zhao, Q., Wortman, J.R., Bidwell, S.L., Alsmark, U.C., Besteiro, S., Sicheritz-Ponten, T., Noel, C.J., Dacks, J.B., Foster, P.G., Simillion, C., Van de Peer, Y., Miranda-Saavedra, D., Barton, G.J., Westrop, G.D., Muller, S., Dessi, D., Fiori, P.L., Ren, Q., Paulsen, I., Zhang, H., Bastida-Corcuera, F.D., Simoes-Barbosa, A., Brown, M.T., Hayes, R.D., Mukherjee, M., Okumura, C.Y., Schneider, R., Smith, A.J., Vanacova, S., Villalvazo, M., Haas, B.J., Pertea, M., Feldblyum, T.V., Utterback, T.R., Shu, C.L., Osoegawa, K., de Jong, P.J., Hrdy, I., Horvathova, L., Zubacova, Z., Dolezal, P., Malik, S.B., Logsdon, J.M., Jr., Henze, K., Gupta, A., Wang, C.C., Dunne, R.L., Upcroft, J.A., Upcroft, P., White, O., Salzberg, S.L., Tang, P., Chiu, C.H., Lee, Y.S., Embley, T.M., Coombs, G.H., Mottram, J.C., Tachezy, J., Fraser-Liggett, C.M., Johnson, P.J., 2007. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. *Science* 315, 207-212.
- Carlton, J.M., Adams, J.H., Silva, J.C., Bidwell, S.L., Lorenzi, H., Caler, E., Crabtree, J., Angiuoli, S.V., Merino, E.F., Amedeo, P., Cheng, Q., Coulson, R.M., Crabb, B.S., Del Portillo, H.A., Essien, K., Feldblyum, T.V., Fernandez-Becerra, C., Gilson, P.R., Gueye, A.H., Guo, X., Kang'a, S., Kooij, T.W., Korsinczky, M., Meyer, E.V., Nene, V., Paulsen, I., White, O., Ralph, S.A., Ren, Q., Sargeant, T.J., Salzberg, S.L., Stoeckert, C.J., Sullivan, S.A., Yamamoto, M.M., Hoffman, S.L., Wortman, J.R., Gardner, M.J., Galinski, M.R., Barnwell, J.W., Fraser-Liggett, C.M., 2008. Comparative genomics of the neglected human malaria parasite *Plasmodium vivax*. *Nature* 455, 757-763.
- Casadevall, A., Pirofski, L., 2001. Host-pathogen interactions: the attributes of virulence. *J Infect Dis* 184, 337-344.
- Cavalier-Smith, T., 1986. The kingdoms Chromista : origin and systematics. *Progress in Physiological Research Biopress Ltd.* 4, 366-369.
- Cavalier-Smith, T., Chao, E.E., 2006. Phylogeny and megasystematics of phagotrophic heterokonts (kingdom Chromista). *J Mol Evol* 62, 388-420.
- Chapman, C.B., Mitchell, G.F., 1982. Proteolytic cleavage of immunoglobulin by enzymes released by *Fasciola hepatica*. *Vet Parasitol* 11, 165-178.
- Chaudhary, K., Roos, D.S., 2005. Protozoan genomics for drug discovery. *Nat Biotechnol* 23, 1089-1091.
- Choi, P., Reiser, H., 1998. IL-4: role in disease and regulation of production. *Clin Exp Immunol* 113, 317-319.
- Clark, C.G., 1997. Extensive genetic diversity in *Blastocystis hominis*. *Mol Biochem Parasitol* 87, 79-83.
- Clark, C.G., Alsmark, U.C., Tazreiter, M., Saito-Nakano, Y., Ali, V., Marion, S., Weber, C., Mukherjee, C., Bruchhaus, I., Tannich, E., Leippe, M., Sicheritz-Ponten, T., Foster, P.G., Samuelson, J., Noel, C.J., Hirt, R.P., Embley, T.M., Gilchrist, C.A., Mann, B.J., Singh, U., Ackers, J.P., Bhattacharya, S., Bhattacharya, A., Lohia, A., Guillen, N., Duchene, M., Nozaki, T., Hall, N., 2007. Structure and content of the *Entamoeba histolytica* genome. *Adv Parasitol* 65, 51-190.
- Coombs, G.H., Goldberg, D.E., Klemba, M., Berry, C., Kay, J., Mottram, J.C., 2001. Aspartic proteases of *Plasmodium falciparum* and other parasitic protozoa as drug targets. *Trends Parasitol* 17, 532-537.
- Cooper, J.B., 2002. Aspartic proteinases in disease: a structural perspective. *Curr Drug Targets* 3, 155-173.

- Davies, D.R., 1990. The structure and function of the aspartic proteinases. *Annu Rev Biophys Biophys Chem* 19, 189-215.
- Denoeud, F., Roussel, M., Noel, B., Wawrzyniak, I., Da Silva, C., Diogon, M., Viscogliosi, E., Brochier-Armanet, C., Couloux, A., Poulain, J., Segurens, B., Anthouard, V., Texier, C., Blot, N., Poirier, P., Ng, G.C., Tan, K.S., Artiguenave, F., Jaillon, O., Aury, J.M., Delbac, F., Wincker, P., Vivares, C.P., El Alaoui, H., 2011. Genome sequence of the stramenopile *Blastocystis*, a human anaerobic parasite. *Genome Biol* 12, R29.
- Dickinson, D.P., 2002. Cysteine peptidases of mammals: their biological roles and potential effects in the oral cavity and other tissues in health and disease. *Crit Rev Oral Biol Med* 13, 238-275.
- Dinareello, C.A., 1999. Interleukin-18. *Methods* 19, 121-132.
- Dixit, A.K., Dixit, P., Sharma, R.L., 2008. Immunodiagnostic/protective role of cathepsin L cysteine proteinases secreted by *Fasciola* species. *Vet Parasitol* 154, 177-184.
- Dogruman-Al, F., Simsek, Z., Boorum, K., Ekici, E., Sahin, M., Tuncer, C., Kustimur, S., Altinbas, A., 2010. Comparison of methods for detection of *Blastocystis* infection in routinely submitted stool samples, and also in IBS/IBD Patients in Ankara, Turkey. *PLoS One* 5, e15484.
- Dominguez-Marquez, M.V., Guna, R., Munoz, C., Gomez-Munoz, M.T., Borrás, R., 2009. High prevalence of subtype 4 among isolates of *Blastocystis hominis* from symptomatic patients of a health district of Valencia (Spain). *Parasitol Res* 105, 949-955.
- Doyle, P.W., Helgason, M.M., Mathias, R.G., Proctor, E.M., 1990. Epidemiology and pathogenicity of *Blastocystis hominis*. *J Clin Microbiol* 28, 116-121.
- Duda, A., Stenzel, D.J., Boreham, P.F., 1998. Detection of *Blastocystis* sp. in domestic dogs and cats. *Vet Parasitol* 76, 9-17.
- Dunn, L.A., Boreham, P.F., Stenzel, D.J., 1989. Ultrastructural variation of *Blastocystis hominis* stocks in culture. *Int J Parasitol* 19, 43-56.
- Durot, M., Bourguignon, P.Y., Schachter, V., 2009. Genome-scale models of bacterial metabolism: reconstruction and applications. *FEMS Microbiol Rev* 33, 164-190.
- El-Sayed, N.M., Myler, P.J., Bartholomeu, D.C., Nilsson, D., Aggarwal, G., Tran, A.N., Ghedin, E., Worthey, E.A., Delcher, A.L., Blandin, G., Westenberger, S.J., Caler, E., Cerqueira, G.C., Branche, C., Haas, B., Anupama, A., Arner, E., Aslund, L., Attipoe, P., Bontempi, E., Bringaud, F., Burton, P., Cadag, E., Campbell, D.A., Carrington, M., Crabtree, J., Darban, H., da Silveira, J.F., de Jong, P., Edwards, K., Englund, P.T., Fazelina, G., Feldblyum, T., Ferella, M., Frasch, A.C., Gull, K., Horn, D., Hou, L., Huang, Y., Kindlund, E., Klingbeil, M., Kluge, S., Koo, H., Lacerda, D., Levin, M.J., Lorenzi, H., Louie, T., Machado, C.R., McCulloch, R., McKenna, A., Mizuno, Y., Mottram, J.C., Nelson, S., Ochaya, S., Osoegawa, K., Pai, G., Parsons, M., Pentony, M., Pettersson, U., Pop, M., Ramirez, J.L., Rinta, J., Robertson, L., Salzberg, S.L., Sanchez, D.O., Seyler, A., Sharma, R., Shetty, J., Simpson, A.J., Sisk, E., Tammi, M.T., Tarleton, R., Teixeira, S., Van Aken, S., Vogt, C., Ward, P.N., Wickstead, B., Wortman, J., White, O., Fraser, C.M., Stuart, K.D., Andersson, B., 2005. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science* 309, 409-415.
- Elwakil, H.S., Hewedi, I.H., 2010. Pathogenic potential of *Blastocystis hominis* in laboratory mice. *Parasitol Res* 107, 685-689.
- Eroglu, F., Koltas, I.S., 2010. Evaluation of the transmission mode of *B. hominis* by using PCR method. *Parasitol Res* 107, 841-845.
- Ertug, S., Dost, T., Ertabaklar, H., Gultekin, B., 2009. [The effect of trimethoprim-sulfamethoxazole in *Blastocystis hominis* infection]. *Turkiye Parazitoloj Derg* 33, 270-272.
- Fabian, I., Kletter, Y., Mor, S., Geller-Bernstein, C., Ben-Yaakov, M., Volovitz, B., Golde, D.W., 1992. Activation of human eosinophil and neutrophil functions by haematopoietic growth factors: comparisons of IL-1, IL-3, IL-5 and GM-CSF. *Br J Haematol* 80, 137-143.
- Fairlamb, A.H., Cerami, A., 1992. Metabolism and functions of trypanothione in the Kinetoplastida. *Annu Rev Microbiol* 46, 695-729.
- Franzen, O., Jerlstrom-Hultqvist, J., Castro, E., Sherwood, E., Ankarklev, J., Reiner, D.S., Palm, D., Andersson, J.O., Andersson, B., Svard, S.G., 2009. Draft genome sequencing of *giardia*

- intestinalis assemblage B isolate GS: is human giardiasis caused by two different species? *PLoS Pathog* 5, e1000560.
- Garavelli, P.L., Scaglione, L., 1989. Blastocystosis. An epidemiological study. *Microbiologica* 12, 349-350.
- Garcia-Nieto, R.M., Rico-Mata, R., Arias-Negrete, S., Avila, E.E., 2008. Degradation of human secretory IgA1 and IgA2 by *Entamoeba histolytica* surface-associated proteolytic activity. *Parasitol Int* 57, 417-423.
- Garcia-Zepeda, E.A., Combadiere, C., Rothenberg, M.E., Sarafi, M.N., Lavigne, F., Hamid, Q., Murphy, P.M., Luster, A.D., 1996. Human monocyte chemoattractant protein (MCP)-4 is a novel CC chemokine with activities on monocytes, eosinophils, and basophils induced in allergic and nonallergic inflammation that signals through the CC chemokine receptors (CCR)-2 and -3. *J Immunol* 157, 5613-5626.
- Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S., Paulsen, I.T., James, K., Eisen, J.A., Rutherford, K., Salzberg, S.L., Craig, A., Kyes, S., Chan, M.S., Nene, V., Shallom, S.J., Suh, B., Peterson, J., Angiuoli, S., Perte, M., Allen, J., Selengut, J., Haft, D., Mather, M.W., Vaidya, A.B., Martin, D.M., Fairlamb, A.H., Fraunholz, M.J., Roos, D.S., Ralph, S.A., McFadden, G.I., Cummings, L.M., Subramanian, G.M., Mungall, C., Venter, J.C., Carucci, D.J., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M., Barrell, B., 2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419, 498-511.
- Gardner, T.B., Hill, D.R., 2001. Treatment of giardiasis. *Clin Microbiol Rev* 14, 114-128.
- Giacometti, A., Cirioni, O., Fiorentini, A., Fortuna, M., Scalise, G., 1999. Irritable bowel syndrome in patients with *Blastocystis hominis* infection. *Eur J Clin Microbiol Infect Dis* 18, 436-439.
- Golden, O., Flynn, R.J., Read, C., Sekiya, M., Donnelly, S.M., Stack, C., Dalton, J.P., Mulcahy, G., 2010. Protection of cattle against a natural infection of *Fasciola hepatica* by vaccination with recombinant cathepsin L1 (rFhCL1). *Vaccine* 28, 5551-5557.
- Greenbaum, D.C., Baruch, A., Grainger, M., Bozdech, Z., Medzihradsky, K.F., Engel, J., DeRisi, J., Holder, A.A., Bogoy, M., 2002. A role for the protease falcipain 1 in host cell invasion by the human malaria parasite. *Science* 298, 2002-2006.
- Grudkowska, M., Zagdanska, B., 2004. Multifunctional role of plant cysteine proteinases. *Acta Biochim Pol* 51, 609-624.
- Gupta, R., Parsi, K., 2006. Chronic urticaria due to *Blastocystis hominis*. *Australas J Dermatol* 47, 117-119.
- Haas, B.J., Kamoun, S., Zody, M.C., Jiang, R.H., Handsaker, R.E., Cano, L.M., Grabherr, M., Kodira, C.D., Raffaele, S., Torto-Alalibo, T., Bozkurt, T.O., Ah-Fong, A.M., Alvarado, L., Anderson, V.L., Armstrong, M.R., Avrova, A., Baxter, L., Beynon, J., Boevink, P.C., Bollmann, S.R., Bos, J.I., Bulone, V., Cai, G., Cakir, C., Carrington, J.C., Chawner, M., Conti, L., Costanzo, S., Ewan, R., Fahlgren, N., Fischbach, M.A., Fugelstad, J., Gilroy, E.M., Gnerre, S., Green, P.J., Grenville-Briggs, L.J., Griffith, J., Grunwald, N.J., Horn, K., Horner, N.R., Hu, C.H., Huitema, E., Jeong, D.H., Jones, A.M., Jones, J.D., Jones, R.W., Karlsson, E.K., Kunjeti, S.G., Lamour, K., Liu, Z., Ma, L., Maclean, D., Chibucos, M.C., McDonald, H., McWalters, J., Meijer, H.J., Morgan, W., Morris, P.F., Munro, C.A., O'Neill, K., Ospina-Giraldo, M., Pinzon, A., Pritchard, L., Ramsahoye, B., Ren, Q., Restrepo, S., Roy, S., Sadanandom, A., Savidor, A., Schornack, S., Schwartz, D.C., Schumann, U.D., Schwessinger, B., Seyer, L., Sharpe, T., Silvar, C., Song, J., Studholme, D.J., Sykes, S., Thines, M., van de Vondervoort, P.J., Phuntumart, V., Wawra, S., Weide, R., Win, J., Young, C., Zhou, S., Fry, W., Meyers, B.C., van West, P., Ristaino, J., Govers, F., Birch, P.R., Whisson, S.C., Judelson, H.S., Nusbaum, C., 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature* 461, 393-398.
- Hameed, D.M., Hassanin, O.M., Zuel-Fakkar, N.M., 2011. Association of *Blastocystis hominis* genetic subtypes with urticaria. *Parasitol Res* 108, 553-560.
- Haresh, K., Suresh, K., Khairul Anus, A., Saminathan, S., 1999. Isolate resistance of *Blastocystis hominis* to metronidazole. *Trop Med Int Health* 4, 274-277.

- Hedstrom, L., 2002. An overview of serine proteases. *Curr Protoc Protein Sci* Chapter 21, Unit 21 10.
- Hernandez, H., Sariego, I., Garber, G., Delgado, R., Lopez, O., Sarracent, J., 2004. Monoclonal antibodies against a 62 kDa proteinase of *Trichomonas vaginalis* decrease parasite cytoadherence to epithelial cells and confer protection in mice. *Parasite Immunol* 26, 119-125.
- Hirata, T., Nakamura, H., Kinjo, N., Hokama, A., Kinjo, F., Yamane, N., Fujita, J., 2007. Prevalence of *Blastocystis hominis* and *Strongyloides stercoralis* infection in Okinawa, Japan. *Parasitol Res* 101, 1717-1719.
- Ho, L.C., Singh, M., Suresh, G., Ng, G.C., Yap, E.H., 1993. Axenic culture of *Blastocystis hominis* in Iscove's modified Dulbecco's medium. *Parasitol Res* 79, 614-616.
- Horimoto, Y., Dee, D.R., Yada, R.Y., 2009. Multifunctional aspartic peptidase prosegments. *N Biotechnol* 25, 318-324.
- Hussein, E.M., Hussein, A.M., Eida, M.M., Atwa, M.M., 2008. Pathophysiological variability of different genotypes of human *Blastocystis hominis* Egyptian isolates in experimentally infected rats. *Parasitol Res* 102, 853-860.
- Iguchi, A., Ebisu, A., Nagata, S., Saitou, Y., Yoshikawa, H., Iwatani, S., Kimata, I., 2007. Infectivity of different genotypes of human *Blastocystis hominis* isolates in chickens and rats. *Parasitol Int* 56, 107-112.
- Iguchi, A., Yoshikawa, H., Yamada, M., Kimata, I., Arizono, N., 2009. Expression of interferon gamma and proinflammatory cytokines in the cecal mucosa of rats experimentally infected with *Blastocystis* sp. strain RN94-9. *Parasitol Res* 105, 135-140.
- Ivens, A.C., Peacock, C.S., Worthey, E.A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M.A., Adlem, E., Aert, R., Anupama, A., Apostolou, Z., Attipoe, P., Bason, N., Bauser, C., Beck, A., Beverley, S.M., Bianchetti, G., Borzym, K., Bothe, G., Bruschi, C.V., Collins, M., Cadag, E., Ciarloni, L., Clayton, C., Coulson, R.M., Cronin, A., Cruz, A.K., Davies, R.M., De Gaudenzi, J., Dobson, D.E., Duesterhoeft, A., Fazelina, G., Fosker, N., Frasch, A.C., Fraser, A., Fuchs, M., Gabel, C., Goble, A., Goffeau, A., Harris, D., Hertz-Fowler, C., Hilbert, H., Horn, D., Huang, Y., Klages, S., Knights, A., Kube, M., Larke, N., Litvin, L., Lord, A., Louie, T., Marra, M., Masuy, D., Matthews, K., Michaeli, S., Mottram, J.C., Muller-Auer, S., Munden, H., Nelson, S., Norbertczak, H., Oliver, K., O'Neil, S., Pentony, M., Pohl, T.M., Price, C., Purnelle, B., Quail, M.A., Rabinowitsch, E., Reinhardt, R., Rieger, M., Rinta, J., Robben, J., Robertson, L., Ruiz, J.C., Rutter, S., Saunders, D., Schafer, M., Schein, J., Schwartz, D.C., Seeger, K., Seyler, A., Sharp, S., Shin, H., Sivam, D., Squares, R., Squares, S., Tosato, V., Vogt, C., Volckaert, G., Wambutt, R., Warren, T., Wedler, H., Woodward, J., Zhou, S., Zimmermann, W., Smith, D.F., Blackwell, J.M., Stuart, K.D., Barrell, B., Myler, P.J., 2005. The genome of the kinetoplastid parasite, *Leishmania major*. *Science* 309, 436-442.
- Jiang, J.B., He, J.G., 1993. Taxonomic status of *Blastocystis hominis*. *Parasitol Today* 9, 2-3.
- Johnson, A.M., Thanou, A., Boreham, P.F., Baverstock, P.R., 1989. *Blastocystis hominis*: phylogenetic affinities determined by rRNA sequence comparison. *Exp Parasitol* 68, 283-288.
- Jones, M.S., 2nd, Ganac, R.D., Hiser, G., Hudson, N.R., Le, A., Whipps, C.M., 2008. Detection of *Blastocystis* from stool samples using real-time PCR. *Parasitol Res* 103, 551-557.
- Katinka, M.D., Duprat, S., Cornillot, E., Metenier, G., Thomarat, F., Prensier, G., Barbe, V., Peyretailade, E., Brottier, P., Wincker, P., Delbac, F., El Alaoui, H., Peyret, P., Saurin, W., Gouy, M., Weissenbach, J., Vivares, C.P., 2001. Genome sequence and gene compaction of the eukaryote parasite *Encephalitozoon cuniculi*. *Nature* 414, 450-453.
- Katsarou-Katsari, A., Vassalos, C.M., Tzanetou, K., Spanakos, G., Papadopoulou, C., Vakalis, N., 2008. Acute urticaria associated with amoeboid forms of *Blastocystis* sp. subtype 3. *Acta Derm Venereol* 88, 80-81.
- Kaur, T., Sobti, R.C., Kaur, S., 2011. Cocktail of gp63 and Hsp70 induces protection against *Leishmania donovani* in BALB/c mice. *Parasite Immunol* 33, 95-103.
- Kaya, S., Cetin, E.S., Aridogan, B.C., Arikan, S., Demirci, M., 2007. Pathogenicity of *Blastocystis hominis*, a clinical reevaluation. *Turkiye Parazitoloj Derg* 31, 184-187.

- Klemba, M., Goldberg, D.E., 2002. Biological roles of proteases in parasitic protozoa. *Annu Rev Biochem* 71, 275-305.
- Kniel, K.E., Sumner, S.S., Pierson, M.D., Zajac, A.M., Hackney, C.R., Fayer, R., Lindsay, D.S., 2004. Effect of hydrogen peroxide and other protease inhibitors on *Cryptosporidium parvum* excystation and in vitro development. *J Parasitol* 90, 885-888.
- Kooij, T.W., Carlton, J.M., Bidwell, S.L., Hall, N., Ramesar, J., Janse, C.J., Waters, A.P., 2005. A *Plasmodium* whole-genome synteny map: indels and synteny breakpoints as foci for species-specific genes. *PLoS Pathog* 1, e44.
- Koussis, K., Withers-Martinez, C., Yeoh, S., Child, M., Hackett, F., Knuepfer, E., Juliano, L., Woehlbier, U., Bujard, H., Blackman, M.J., 2009. A multifunctional serine protease primes the malaria parasite for red blood cell invasion. *Embo J* 28, 725-735.
- Lee, B.J., Bak, Y.T., 2011. Irritable bowel syndrome, gut microbiota and probiotics. *J Neurogastroenterol Motil* 17, 252-266.
- Leelayoova, S., Taamasri, P., Rangsin, R., Naaglor, T., Thathaisong, U., Mungthin, M., 2002. In-vitro cultivation: a sensitive method for detecting *Blastocystis hominis*. *Ann Trop Med Parasitol* 96, 803-807.
- Leelayoova, S., Rangsin, R., Taamasri, P., Naaglor, T., Thathaisong, U., Mungthin, M., 2004. Evidence of waterborne transmission of *Blastocystis hominis*. *Am J Trop Med Hyg* 70, 658-662.
- Leelayoova, S., Siripattanapipong, S., Thathaisong, U., Naaglor, T., Taamasri, P., Piyaraj, P., Mungthin, M., 2008. Drinking water: a possible source of *Blastocystis* spp. subtype 1 infection in schoolchildren of a rural community in central Thailand. *Am J Trop Med Hyg* 79, 401-406.
- Leelayoova, S., Siripattanapipong, S., Naaglor, T., Taamasri, P., Mungthin, M., 2009. Prevalence of intestinal parasitic infections in military personnel and military dogs, Thailand. *J Med Assoc Thai* 92 Suppl 1, S53-59.
- Lipscomb, W.N., Strater, N., 1996. Recent Advances in Zinc Enzymology. *Chem Rev* 96, 2375-2434.
- Loftus, B., Anderson, I., Davies, R., Alsmark, U.C., Samuelson, J., Amedeo, P., Roncaglia, P., Berriman, M., Hirt, R.P., Mann, B.J., Nozaki, T., Suh, B., Pop, M., Duchene, M., Ackers, J., Tannich, E., Leippe, M., Hofer, M., Bruchhaus, I., Willhoeft, U., Bhattacharya, A., Chillingworth, T., Churcher, C., Hance, Z., Harris, B., Harris, D., Jagels, K., Moule, S., Mungall, K., Ormond, D., Squares, R., Whitehead, S., Quail, M.A., Rabinowitsch, E., Norbertczak, H., Price, C., Wang, Z., Guillen, N., Gilchrist, C., Stroup, S.E., Bhattacharya, S., Lohia, A., Foster, P.G., Sicheritz-Ponten, T., Weber, C., Singh, U., Mukherjee, C., El-Sayed, N.M., Petri, W.A., Jr., Clark, C.G., Embley, T.M., Barrell, B., Fraser, C.M., Hall, N., 2005. The genome of the protist parasite *Entamoeba histolytica*. *Nature* 433, 865-868.
- Long, H.Y., Handschack, A., Konig, W., Ambrosch, A., 2001. *Blastocystis hominis* modulates immune responses and cytokine release in colonic epithelial cells. *Parasitol Res* 87, 1029-1030.
- Lorenzi, H.A., Puiu, D., Miller, J.R., Brinkac, L.M., Amedeo, P., Hall, N., Caler, E.V., 2010. New assembly, reannotation and analysis of the *Entamoeba histolytica* genome reveal new genomic features and protein content information. *PLoS Negl Trop Dis* 4, e716.
- Lotter, H., Tannich, E., 2006. The current status of an amebiasis vaccine. *Arch Med Res* 37, 292-296.
- Luster, A.D., 1998. Chemokines--chemotactic cytokines that mediate inflammation. *N Engl J Med* 338, 436-445.
- MacFarlane, R.C., Singh, U., 2006. Identification of differentially expressed genes in virulent and nonvirulent *Entamoeba* species: potential implications for amebic pathogenesis. *Infect Immun* 74, 340-351.
- Malemud, C.J., 2006. Matrix metalloproteinases (MMPs) in health and disease: an overview. *Front Biosci* 11, 1696-1701.
- McDonald, J.K., 1985. An overview of protease specificity and catalytic mechanisms: aspects related to nomenclature and classification. *Histochem J* 17, 773-785.
- McGwire, B.S., Chang, K.P., Engman, D.M., 2003. Migration through the extracellular matrix by the parasitic protozoan *Leishmania* is enhanced by surface metalloprotease gp63. *Infect Immun* 71, 1008-1010.

- McKerrow, J.H., Sun, E., Rosenthal, P.J., Bouvier, J., 1993. The proteases and pathogenicity of parasitic protozoa. *Annu Rev Microbiol* 47, 821-853.
- McKerrow, J.H., Rosenthal, P.J., Swenerton, R., Doyle, P., 2008. Development of protease inhibitors for protozoan infections. *Curr Opin Infect Dis* 21, 668-672.
- Mehlhorn, H., 1988. *Blastocystis hominis*, Brumpt 1912: are there different stages or species? *Parasitol Res* 74, 393-395.
- Mendoza-Lopez, M.R., Becerril-Garcia, C., Fattel-Facenda, L.V., Avila-Gonzalez, L., Ruiz-Tachiquin, M.E., Ortega-Lopez, J., Arroyo, R., 2000. CP30, a cysteine proteinase involved in *Trichomonas vaginalis* cytoadherence. *Infect Immun* 68, 4907-4912.
- Micheloud, D., Jensen, J., Fernandez-Cruz, E., Carbone, J., 2007. [Chronic angioedema and *blastocystis hominis* infection]. *Rev Gastroenterol Peru* 27, 191-193.
- Mirza, H., Tan, K.S., 2009. *Blastocystis* exhibits inter- and intra-subtype variation in cysteine protease activity. *Parasitol Res* 104, 355-361.
- Mirza, H., Teo, J.D., Upcroft, J., Tan, K.S., 2011. A rapid, high-throughput viability assay for *Blastocystis* spp. reveals metronidazole resistance and extensive subtype-dependent variations in drug susceptibilities. *Antimicrob Agents Chemother* 55, 637-648.
- Moe, K.T., Singh, M., Howe, J., Ho, L.C., Tan, S.W., Ng, G.C., Chen, X.Q., Yap, E.H., 1996. Observations on the ultrastructure and viability of the cystic stage of *Blastocystis hominis* from human feces. *Parasitol Res* 82, 439-444.
- Moe, K.T., Singh, M., Howe, J., Ho, L.C., Tan, S.W., Chen, X.Q., Ng, G.C., Yap, E.H., 1997. Experimental *Blastocystis hominis* infection in laboratory mice. *Parasitol Res* 83, 319-325.
- Moe, K.T., Singh, M., Gopalakrishnakone, P., Ho, L.C., Tan, S.W., Chen, X.Q., Yap, E.H., 1998. Cytopathic effect of *Blastocystis hominis* after intramuscular inoculation into laboratory mice. *Parasitol Res* 84, 450-454.
- Moghaddam, D.D., Ghadirian, E., Azami, M., 2005. *Blastocystis hominis* and the evaluation of efficacy of metronidazole and trimethoprim/sulfamethoxazole. *Parasitol Res* 96, 273-275.
- Mohsen, A., Hossein, H., 2009. Gastrointestinal parasites of stray cats in Kashan, Iran. *Trop Biomed* 26, 16-22.
- Moncada, D., Keller, K., Chadee, K., 2003. *Entamoeba histolytica* cysteine proteinases disrupt the polymeric structure of colonic mucin and alter its protective function. *Infect Immun* 71, 838-844.
- Morrison, H.G., McArthur, A.G., Gillin, F.D., Aley, S.B., Adam, R.D., Olsen, G.J., Best, A.A., Cande, W.Z., Chen, F., Cipriano, M.J., Davids, B.J., Dawson, S.C., Elmendorf, H.G., Hehl, A.B., Holder, M.E., Huse, S.M., Kim, U.U., Lasek-Nesselquist, E., Manning, G., Nigam, A., Nixon, J.E., Palm, D., Passamaneck, N.E., Prabhu, A., Reich, C.I., Reiner, D.S., Samuelson, J., Svard, S.G., Sogin, M.L., 2007. Genomic minimalism in the early diverging intestinal parasite *Giardia lamblia*. *Science* 317, 1921-1926.
- Mottram, J.C., Brooks, D.R., Coombs, G.H., 1998. Roles of cysteine proteinases of trypanosomes and *Leishmania* in host-parasite interactions. *Curr Opin Microbiol* 1, 455-460.
- Nakamura, Y., Hashimoto, T., Yoshikawa, H., Kamaishi, T., Nakamura, F., Okamoto, K., Hasegawa, M., 1996. Phylogenetic position of *Blastocystis hominis* that contains cytochrome-free mitochondria, inferred from the protein phylogeny of elongation factor 1 alpha. *Mol Biochem Parasitol* 77, 241-245.
- Nasirudeen, A.M., Tan, K.S., 2004. Isolation and characterization of the mitochondrion-like organelle from *Blastocystis hominis*. *J Microbiol Methods* 58, 101-109.
- Nigro, L., Larocca, L., Massarelli, L., Patamia, I., Minniti, S., Palermo, F., Cacopardo, B., 2003. A placebo-controlled treatment trial of *Blastocystis hominis* infection with metronidazole. *J Travel Med* 10, 128-130.
- Nimri, L., Batchoun, R., 1994. Intestinal colonization of symptomatic and asymptomatic schoolchildren with *Blastocystis hominis*. *J Clin Microbiol* 32, 2865-2866.
- Noel, C., Dufernez, F., Gerbod, D., Edgcomb, V.P., Delgado-Viscogliosi, P., Ho, L.C., Singh, M., Wintjens, R., Sogin, M.L., Capron, M., Pierce, R., Zenner, L., Viscogliosi, E., 2005. Molecular

- phylogenies of *Blastocystis* isolates from different hosts: implications for genetic diversity, identification of species, and zoonosis. *J Clin Microbiol* 43, 348-355.
- North, M.J., 1994. Cysteine endopeptidases of parasitic protozoa. *Methods Enzymol* 244, 523-539.
- Obermajer, N., Jevnikar, Z., Doljak, B., Kos, J., 2008. Role of cysteine cathepsins in matrix degradation and cell signalling. *Connect Tissue Res* 49, 193-196.
- Ocadiz, R., Orozco, E., Carrillo, E., Quintas, L.I., Ortega-Lopez, J., Garcia-Perez, R.M., Sanchez, T., Castillo-Juarez, B.A., Garcia-Rivera, G., Rodriguez, M.A., 2005. EhCP112 is an *Entamoeba histolytica* secreted cysteine protease that may be involved in the parasite-virulence. *Cell Microbiol* 7, 221-232.
- Ok, U.Z., Girginkardesler, N., Balcioglu, C., Ertan, P., Pirildar, T., Kilimcioglu, A.A., 1999. Effect of trimethoprim-sulfamethaxazole in *Blastocystis hominis* infection. *Am J Gastroenterol* 94, 3245-3247.
- Olivos-Garcia, A., Tello, E., Nequiz-Avendano, M., Gonzalez-Canto, A., Lopez-Vancell, R., Garcia de Leon, M.C., Montfort, I., Perez-Tamayo, R., 2004. Cysteine proteinase activity is required for survival of the parasite in experimental acute amoebic liver abscesses in hamsters. *Parasitology* 129, 19-25.
- Ozyurt, M., Kurt, O., Molbak, K., Nielsen, H.V., Haznedaroglu, T., Stensvold, C.R., 2008. Molecular epidemiology of *Blastocystis* infections in Turkey. *Parasitol Int* 57, 300-306.
- Page, M.J., Di Cera, E., 2008. Serine peptidases: classification, structure and function. *Cell Mol Life Sci* 65, 1220-1236.
- Parkar, U., Traub, R.J., Kumar, S., Mungthin, M., Vitali, S., Leelayoova, S., Morris, K., Thompson, R.C., 2007. Direct characterization of *Blastocystis* from faeces by PCR and evidence of zoonotic potential. *Parasitology* 134, 359-367.
- Parkar, U., Traub, R.J., Vitali, S., Elliot, A., Levecke, B., Robertson, I., Geurden, T., Steele, J., Drake, B., Thompson, R.C., 2010. Molecular characterization of *Blastocystis* isolates from zoo animals and their animal-keepers. *Vet Parasitol* 169, 8-17.
- Parkes, G.C., Brostoff, J., Whelan, K., Sanderson, J.D., 2008. Gastrointestinal microbiota in irritable bowel syndrome: their role in its pathogenesis and treatment. *Am J Gastroenterol* 103, 1557-1567.
- Peacock, C.S., Seeger, K., Harris, D., Murphy, L., Ruiz, J.C., Quail, M.A., Peters, N., Adlem, E., Tivey, A., Aslett, M., Kerhornou, A., Ivens, A., Fraser, A., Rajandream, M.A., Carver, T., Norbertczak, H., Chillingworth, T., Hance, Z., Jagels, K., Moule, S., Ormond, D., Rutter, S., Squares, R., Whitehead, S., Rabbinowitsch, E., Arrowsmith, C., White, B., Thurston, S., Bringaud, F., Baldauf, S.L., Faulconbridge, A., Jeffares, D., Depledge, D.P., Oyola, S.O., Hilley, J.D., Brito, L.O., Tosi, L.R., Barrell, B., Cruz, A.K., Mottram, J.C., Smith, D.F., Berriman, M., 2007. Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. *Nat Genet* 39, 839-847.
- Pegelow, K., Gross, R., Pietrzik, K., Lukito, W., Richards, A.L., Fryauff, D.J., 1997. Parasitological and nutritional situation of school children in the Sukaraja district, West Java, Indonesia. *Southeast Asian J Trop Med Public Health* 28, 173-190.
- Pertuz Belloso, S., Ostoa Saloma, P., Benitez, I., Soldevila, G., Olivos, A., Garcia-Zepeda, E., 2004. *Entamoeba histolytica* cysteine protease 2 (EhCP2) modulates leucocyte migration by proteolytic cleavage of chemokines. *Parasite Immunol* 26, 237-241.
- Piche, T., Barbara, G., Aubert, P., Bruley des Varannes, S., Dainese, R., Nano, J.L., Cremon, C., Stanghellini, V., De Giorgio, R., Galmiche, J.P., Neunlist, M., 2009. Impaired intestinal barrier integrity in the colon of patients with irritable bowel syndrome: involvement of soluble mediators. *Gut* 58, 196-201.
- Pinney, J.W., Papp, B., Hyland, C., Wambua, L., Westhead, D.R., McConkey, G.A., 2007. Metabolic reconstruction and analysis for parasite genomes. *Trends Parasitol* 23, 548-554.
- Pleass, R.J., Kusel, J.R., Woof, J.M., 2000. Cleavage of human IgE mediated by *Schistosoma mansoni*. *Int Arch Allergy Immunol* 121, 194-204.

- Poirier, P., Wawrzyniak, I., Albert, A., El Alaoui, H., Delbac, F., Livrelli, V., 2011. Development and evaluation of a real-time PCR assay for detection and quantification of *Blastocystis* in human stool samples: a prospective study in patients with hematological malignancies. *J Clin Microbiol*.
- Poulin, R., Combes, C., 1999. The concept of virulence: interpretations and implications. *Parasitol Today* 15, 474-475.
- Puthia, M.K., Vaithilingam, A., Lu, J., Tan, K.S., 2005. Degradation of human secretory immunoglobulin A by *Blastocystis*. *Parasitol Res* 97, 386-389.
- Puthia, M.K., Sio, S.W., Lu, J., Tan, K.S., 2006. *Blastocystis ratti* induces contact-independent apoptosis, F-actin rearrangement, and barrier function disruption in IEC-6 cells. *Infect Immun* 74, 4114-4123.
- Puthia, M.K., Lu, J., Tan, K.S., 2008. *Blastocystis ratti* contains cysteine proteases that mediate interleukin-8 response from human intestinal epithelial cells in an NF-kappaB-dependent manner. *Eukaryot Cell* 7, 435-443.
- Que, X., Reed, S.L., 2000. Cysteine proteinases and the pathogenesis of amebiasis. *Clin Microbiol Rev* 13, 196-206.
- Que, X., Kim, S.H., Sajid, M., Eckmann, L., Dinarello, C.A., McKerrow, J.H., Reed, S.L., 2003. A surface amebic cysteine proteinase inactivates interleukin-18. *Infect Immun* 71, 1274-1280.
- Qutob, D., Hraber, P.T., Sobral, B.W., Gijzen, M., 2000. Comparative analysis of expressed sequences in *Phytophthora sojae*. *Plant Physiol* 123, 243-254.
- Randall, T.A., Dwyer, R.A., Huitema, E., Beyer, K., Cvitanich, C., Kelkar, H., Fong, A.M., Gates, K., Roberts, S., Yatzkan, E., Gaffney, T., Law, M., Testa, A., Torto-Alalibo, T., Zhang, M., Zheng, L., Mueller, E., Windass, J., Binder, A., Birch, P.R., Gisi, U., Govers, F., Gow, N.A., Mauch, F., van West, P., Waugh, M.E., Yu, J., Boller, T., Kamoun, S., Lam, S.T., Judelson, H.S., 2005. Large-scale gene discovery in the oomycete *Phytophthora infestans* reveals likely components of phytopathogenicity shared with true fungi. *Mol Plant Microbe Interact* 18, 229-243.
- Rao, M.B., Tanksale, A.M., Ghatge, M.S., Deshpande, V.V., 1998. Molecular and biotechnological aspects of microbial proteases. *Microbiol Mol Biol Rev* 62, 597-635.
- Rawlings, N.D., Barrett, A.J., 1995. Evolutionary families of metallopeptidases. *Methods Enzymol* 248, 183-228.
- Rawlings, N.D., Barrett, A.J., F, J.J., 2004. introduction : metallopeptidases and their clans *Handbook of Proteolytic Enzymes - 2 Vol Set*, 2nd Ed 1, 231-268.
- Rawlings, N.D., Morton, F.R., Barrett, A.J., 2006. MEROPS: the peptidase database. *Nucleic Acids Res* 34, D270-272.
- Rawlings, N.D., Morton, F.R., Kok, C.Y., Kong, J., Barrett, A.J., 2008. MEROPS: the peptidase database. *Nucleic Acids Res* 36, D320-325.
- Rawlings, N.D., Barrett, A.J., Bateman, A., 2010. MEROPS: the peptidase database. *Nucleic Acids Res* 38, D227-233.
- Rayan, H.Z., Ismail, O.A., El Gayar, E.K., 2007. Prevalence and clinical features of *Dientamoeba fragilis* infections in patients suspected to have intestinal parasitic infection. *J Egypt Soc Parasitol* 37, 599-608.
- Rees-Roberts, D., Mullen, L.M., Gounaris, K., Selkirk, M.E., 2010. Inactivation of the complement anaphylatoxin C5a by secreted products of parasitic nematodes. *Int J Parasitol* 40, 527-532.
- Riisberg, I., Orr, R.J., Kluge, R., Shalchian-Tabrizi, K., Bowers, H.A., Patil, V., Edvardsen, B., Jakobsen, K.S., 2009. Seven gene phylogeny of heterokonts. *Protist* 160, 191-204.
- Rogier, C., Moukoko, C.E., Orlandi-Pradines, E., Briolant, S., 2007. [Antiparasitic vaccines: where are we now?]. *Rev Prat* 57, 183-188.
- Rosenthal, P.J., 2004. Cysteine proteases of malaria parasites. *Int J Parasitol* 34, 1489-1499.
- Rudenskaya, G.N., Pupov, D.V., 2008. Cysteine proteinases of microorganisms and viruses. *Biochemistry (Mosc)* 73, 1-13.
- Sajid, M., McKerrow, J.H., 2002. Cysteine proteases of parasitic organisms. *Mol Biochem Parasitol* 120, 1-21.

- Saksirisampant, W., Prownebon, J., Kulkumthorn, M., Yenthakam, S., Janpla, S., Nuchprayoon, S., 2006. Prevalence of intestinal parasitic infections among school children in the central region of Thailand. *J Med Assoc Thai* 89, 1928-1933.
- Samuelson, J., 1999. Why metronidazole is active against both bacteria and parasites. *Antimicrob Agents Chemother* 43, 1533-1541.
- Santin, M., Gomez-Munoz, M.T., Solano-Aguilar, G., Fayer, R., 2011. Development of a new PCR protocol to detect and subtype *Blastocystis* spp. from humans and animals. *Parasitol Res*.
- Sateriale, A., Huston, C.D., 2011. A Sequential Model of Host Cell Killing and Phagocytosis by *Entamoeba histolytica*. *J Parasitol Res* 2011, 926706.
- Satoskar, A., Bluethmann, H., Alexander, J., 1995. Disruption of the murine interleukin-4 gene inhibits disease progression during *Leishmania mexicana* infection but does not increase control of *Leishmania donovani* infection. *Infect Immun* 63, 4894-4899.
- Schaechter, M., Ingraham, J.L., 2002. What limits genomics, proteomics, transcriptomics? *Int Microbiol* 5, 51-52.
- Scicluna, S.M., Tawari, B., Clark, C.G., 2006. DNA barcoding of *blastocystis*. *Protist* 157, 77-85.
- Shah, P.H., MacFarlane, R.C., Bhattacharya, D., Matese, J.C., Demeter, J., Stroup, S.E., Singh, U., 2005. Comparative genomic hybridizations of *Entamoeba* strains reveal unique genetic fingerprints that correlate with virulence. *Eukaryot Cell* 4, 504-515.
- Silberman, J.D., Sogin, M.L., Leipe, D.D., Clark, C.G., 1996. Human parasite finds taxonomic home. *Nature* 380, 398.
- Singh, M., Suresh, K., Ho, L.C., Ng, G.C., Yap, E.H., 1995. Elucidation of the life cycle of the intestinal protozoan *Blastocystis hominis*. *Parasitol Res* 81, 446-450.
- Sio, S.W., Puthia, M.K., Lee, A.S., Lu, J., Tan, K.S., 2006. Protease activity of *Blastocystis hominis*. *Parasitol Res* 99, 126-130.
- Smooker, P.M., Jayaraj, R., Pike, R.N., Spithill, T.W., 2010. Cathepsin B proteases of flukes: the key to facilitating parasite control? *Trends Parasitol* 26, 506-514.
- Souppart, L., Sancier, G., Cian, A., Wawrzyniak, I., Delbac, F., Capron, M., Dei-Cas, E., Boorom, K., Delhaes, L., Viscogliosi, E., 2009. Molecular epidemiology of human *Blastocystis* isolates in France. *Parasitol Res* 105, 413-421.
- Souppart, L., Moussa, H., Cian, A., Sancier, G., Poirier, P., El Alaoui, H., Delbac, F., Boorom, K., Delhaes, L., Dei-Cas, E., Viscogliosi, E., 2010. Subtype analysis of *Blastocystis* isolates from symptomatic patients in Egypt. *Parasitol Res* 106, 505-511.
- Stanley, S.L., Jr., Zhang, T., Rubin, D., Li, E., 1995. Role of the *Entamoeba histolytica* cysteine proteinase in amebic liver abscess formation in severe combined immunodeficient mice. *Infect Immun* 63, 1587-1590.
- Stensvold, C.R., Arendrup, M.C., Jespersgaard, C., Molbak, K., Nielsen, H.V., 2007a. Detecting *Blastocystis* using parasitologic and DNA-based methods: a comparative study. *Diagn Microbiol Infect Dis* 59, 303-307.
- Stensvold, C.R., Suresh, G.K., Tan, K.S., Thompson, R.C., Traub, R.J., Viscogliosi, E., Yoshikawa, H., Clark, C.G., 2007b. Terminology for *Blastocystis* subtypes--a consensus. *Trends Parasitol* 23, 93-96.
- Stensvold, C.R., Arendrup, M.C., Nielsen, H.V., Bada, A., Thorsen, S., 2008. Symptomatic infection with *Blastocystis* sp. subtype 8 successfully treated with trimethoprim-sulfamethoxazole. *Ann Trop Med Parasitol* 102, 271-274.
- Stensvold, C.R., Alfellani, M.A., Norkov-Lauritsen, S., Prip, K., Victory, E.L., Maddox, C., Nielsen, H.V., Clark, C.G., 2009. Subtype distribution of *Blastocystis* isolates from synanthropic and zoo animals and identification of a new subtype. *Int J Parasitol* 39, 473-479.
- Stensvold, C.R., Smith, H.V., Nagel, R., Olsen, K.E., Traub, R.J., 2010. Eradication of *Blastocystis* carriage with antimicrobials: reality or delusion? *J Clin Gastroenterol* 44, 85-90.
- Stensvold, R., Brillowska-Dabrowska, A., Nielsen, H.V., Arendrup, M.C., 2006. Detection of *Blastocystis hominis* in unpreserved stool specimens by using polymerase chain reaction. *J Parasitol* 92, 1081-1087.

- Stenzel, D.J., Boreham, P.F., 1991. A cyst-like stage of *Blastocystis hominis*. *Int J Parasitol* 21, 613-615.
- Stenzel, D.J., Boreham, P.F., 1996. *Blastocystis hominis* revisited. *Clin Microbiol Rev* 9, 563-584.
- Sun, C.H., McCaffery, J.M., Reiner, D.S., Gillin, F.D., 2003. Mining the *Giardia lamblia* genome for new cyst wall proteins. *J Biol Chem* 278, 21701-21708.
- Suresh, K., Ng, G.C., Ramachandran, N.P., Ho, L.C., Yap, E.H., Singh, M., 1993. In vitro encystment and experimental infections of *Blastocystis hominis*. *Parasitol Res* 79, 456-460.
- Suresh, K., Smith, H., 2004. Comparison of methods for detecting *Blastocystis hominis*. *Eur J Clin Microbiol Infect Dis* 23, 509-511.
- Suresh, K., Venilla, G.D., Tan, T.C., Rohela, M., 2009. In vivo encystation of *Blastocystis hominis*. *Parasitol Res* 104, 1373-1380.
- Surolia, N., Surolia, A., 2001. Triclosan offers protection against blood stages of malaria by inhibiting enoyl-ACP reductase of *Plasmodium falciparum*. *Nat Med* 7, 167-173.
- Tan, K.S., Ng, G.C., Quek, E., Howe, J., Ramachandran, N.P., Yap, E.H., Singh, M., 2000. *Blastocystis hominis*: A simplified, high-efficiency method for clonal growth on solid agar. *Exp Parasitol* 96, 9-15.
- Tan, K.S., Howe, J., Yap, E.H., Singh, M., 2001. Do *Blastocystis hominis* colony forms undergo programmed cell death? *Parasitol Res* 87, 362-367.
- Tan, K.S., Singh, M., Yap, E.H., 2002. Recent advances in *Blastocystis hominis* research: hot spots in terra incognita. *Int J Parasitol* 32, 789-804.
- Tan, K.S., 2004. *Blastocystis* in humans and animals: new insights using modern methodologies. *Vet Parasitol* 126, 121-144.
- Tan, K.S., 2008. New insights on classification, identification, and clinical relevance of *Blastocystis* spp. *Clin Microbiol Rev* 21, 639-665.
- Tan, K.S., Mirza, H., Teo, J.D., Wu, B., Macary, P.A., 2010. Current Views on the Clinical Relevance of *Blastocystis* spp. *Curr Infect Dis Rep* 12, 28-35.
- Tan, S.W., Singh, M., Thong, K.T., Ho, L.C., Moe, K.T., Chen, X.Q., Ng, G.C., Yap, E.H., 1996. Clonal growth of *Blastocystis hominis* in soft agar with sodium thioglycollate. *Parasitol Res* 82, 737-739.
- Tan, T.C., Suresh, K.G., 2006a. Amoeboid form of *Blastocystis hominis* - a detailed ultrastructural insight. *Parasitol Res* 99, 737-742.
- Tan, T.C., Suresh, K.G., 2006b. Predominance of amoeboid forms of *Blastocystis hominis* in isolates from symptomatic patients. *Parasitol Res* 98, 189-193.
- Tan, T.C., Suresh, K.G., 2007. Evidence of plasmotomy in *Blastocystis hominis*. *Parasitol Res* 101, 1521-1525.
- Termmathurapoj, S., Leelayoova, S., Aimpun, P., Thathaisong, U., Nimmanon, T., Taamasri, P., Mungthin, M., 2004. The usefulness of short-term in vitro cultivation for the detection and molecular study of *Blastocystis hominis* in stool specimens. *Parasitol Res* 93, 445-447.
- Tielens, A.G., Rotte, C., van Hellemond, J.J., Martin, W., 2002. Mitochondria as we don't know them. *Trends Biochem Sci* 27, 564-572.
- Torto, T.A., Li, S., Styer, A., Huitema, E., Testa, A., Gow, N.A., van West, P., Kamoun, S., 2003. EST mining and functional expression assays identify extracellular effector proteins from the plant pathogen *Phytophthora*. *Genome Res* 13, 1675-1685.
- Tyler, B.M., Tripathy, S., Zhang, X., Dehal, P., Jiang, R.H., Aerts, A., Arredondo, F.D., Baxter, L., Bensasson, D., Beynon, J.L., Chapman, J., Damasceno, C.M., Dorrance, A.E., Dou, D., Dickerman, A.W., Dubchak, I.L., Garbelotto, M., Gijzen, M., Gordon, S.G., Govers, F., Grunwald, N.J., Huang, W., Ivors, K.L., Jones, R.W., Kamoun, S., Krampis, K., Lamour, K.H., Lee, M.K., McDonald, W.H., Medina, M., Meijer, H.J., Nordberg, E.K., Maclean, D.J., Ospina-Giraldo, M.D., Morris, P.F., Phuntumart, V., Putnam, N.H., Rash, S., Rose, J.K., Sakihama, Y., Salamov, A.A., Savidor, A., Scheuring, C.F., Smith, B.M., Sobral, B.W., Terry, A., Torto-Alalibo, T.A., Win, J., Xu, Z., Zhang, H., Grigoriev, I.V., Rokhsar, D.S., Boore, J.L., 2006. *Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* 313, 1261-1266.

- Variyam, E.P., 2007. Luminal host-defense mechanisms against invasive amebiasis. *Trends Parasitol* 23, 108-111.
- Villena, I., Aubert, D., Pinon, J.M., 2006. [Cell-host-parasite interactions: biodiversity, pathogenesis, environment]. *Ann Pharm Fr* 64, 115-120.
- Vogelberg, C., Stensvold, C.R., Monecke, S., Ditzen, A., Stopsack, K., Heinrich-Grafe, U., Pohlmann, C., 2010. Blastocystis sp. subtype 2 detection during recurrence of gastrointestinal and urticarial symptoms. *Parasitol Int* 59, 469-471.
- Waller, R.F., Keeling, P.J., Donald, R.G., Striepen, B., Handman, E., Lang-Unnasch, N., Cowman, A.F., Besra, G.S., Roos, D.S., McFadden, G.I., 1998. Nuclear-encoded proteins target to the plastid in *Toxoplasma gondii* and *Plasmodium falciparum*. *Proc Natl Acad Sci U S A* 95, 12352-12357.
- Waller, R.F., Ralph, S.A., Reed, M.B., Su, V., Douglas, J.D., Minnikin, D.E., Cowman, A.F., Besra, G.S., McFadden, G.I., 2003. A type II pathway for fatty acid biosynthesis presents drug targets in *Plasmodium falciparum*. *Antimicrob Agents Chemother* 47, 297-301.
- Wang, Z., Samuelson, J., Clark, C.G., Eichinger, D., Paul, J., Van Dellen, K., Hall, N., Anderson, I., Loftus, B., 2003. Gene discovery in the *Entamoeba invadens* genome. *Mol Biochem Parasitol* 129, 23-31.
- Wawrzyniak, I., Roussel, M., Diogon, M., Couloux, A., Texier, C., Tan, K.S., Vivares, C.P., Delbac, F., Wincker, P., El Alaoui, H., 2008. Complete circular DNA in the mitochondria-like organelles of *Blastocystis hominis*. *Int J Parasitol* 38, 1377-1382.
- Wiederanders, B., Kaulmann, G., Schilling, K., 2003. Functions of propeptide parts in cysteine proteases. *Curr Protein Pept Sci* 4, 309-326.
- Wong, K.H., Ng, G.C., Lin, R.T., Yoshikawa, H., Taylor, M.B., Tan, K.S., 2008. Predominance of subtype 3 among *Blastocystis* isolates from a major hospital in Singapore. *Parasitol Res* 102, 663-670.
- Woof, J.M., Kerr, M.A., 2006. The function of immunoglobulin A in immunity. *J Pathol* 208, 270-282.
- Wu, B., Yin, J., Texier, C., Roussel, M., Tan, K.S., 2010. Blastocystis legumain is localized on the cell surface, and specific inhibition of its activity implicates a pro-survival role for the enzyme. *J Biol Chem* 285, 1790-1798.
- Wu, J.W., Chen, X.L., 2011. Extracellular metalloproteases from bacteria. *Appl Microbiol Biotechnol*.
- Ximenez, C., Moran, P., Rojas, L., Valadez, A., Gomez, A., 2009. Reassessment of the epidemiology of amebiasis: state of the art. *Infect Genet Evol* 9, 1023-1032.
- Xu, P., Widmer, G., Wang, Y., Ozaki, L.S., Alves, J.M., Serrano, M.G., Puiu, D., Manque, P., Akiyoshi, D., Mackey, A.J., Pearson, W.R., Dear, P.H., Bankier, A.T., Peterson, D.L., Abrahamsen, M.S., Kapur, V., Tzipori, S., Buck, G.A., 2004. The genome of *Cryptosporidium hominis*. *Nature* 431, 1107-1112.
- Yakoob, J., Jafri, W., Jafri, N., Khan, R., Islam, M., Beg, M.A., Zaman, V., 2004. Irritable bowel syndrome: in search of an etiology: role of *Blastocystis hominis*. *Am J Trop Med Hyg* 70, 383-385.
- Yakoob, J., Jafri, W., Beg, M.A., Abbas, Z., Naz, S., Islam, M., Khan, R., 2010a. Irritable bowel syndrome: is it associated with genotypes of *Blastocystis hominis*. *Parasitol Res* 106, 1033-1038.
- Yakoob, J., Jafri, W., Beg, M.A., Abbas, Z., Naz, S., Islam, M., Khan, R., 2010b. *Blastocystis hominis* and *Dientamoeba fragilis* in patients fulfilling irritable bowel syndrome criteria. *Parasitol Res* 107, 679-684.
- Yao, C., 2010. Major surface protease of trypanosomatids: one size fits all? *Infect Immun* 78, 22-31.
- Yoshikawa, H., Kuwayama, N., Enose, Y., 1995a. Histochemical detection of carbohydrates of *Blastocystis hominis*. *J Eukaryot Microbiol* 42, 70-74.
- Yoshikawa, H., Satoh, J., Enose, Y., 1995b. Light and electron microscopic localization of lipids in *Blastocystis hominis*. *J Electron Microsc (Tokyo)* 44, 100-103.
- Yoshikawa, H., Nagano, I., Wu, Z., Yap, E.H., Singh, M., Takahashi, Y., 1998. Genomic polymorphism among *Blastocystis hominis* strains and development of subtype-specific diagnostic primers. *Mol Cell Probes* 12, 153-159.

- Yoshikawa, H., Abe, N., Iwasawa, M., Kitano, S., Nagano, I., Wu, Z., Takahashi, Y., 2000. Genomic analysis of *Blastocystis hominis* strains isolated from two long-term health care facilities. *J Clin Microbiol* 38, 1324-1330.
- Yoshikawa, H., Abe, N., Wu, Z., 2004a. PCR-based identification of zoonotic isolates of *Blastocystis* from mammals and birds. *Microbiology* 150, 1147-1151.
- Yoshikawa, H., Yoshida, K., Nakajima, A., Yamanari, K., Iwatani, S., Kimata, I., 2004b. Fecal-oral transmission of the cyst form of *Blastocystis hominis* in rats. *Parasitol Res* 94, 391-396.
- Zaman, V., Khan, K.Z., 1994. A concentration technique for obtaining viable cysts of *blastocystis hominis* from faeces. *J Pak Med Assoc* 44, 220-221.
- Zaman, V., Howe, J., Ng, M., 1997. Observations on the surface coat of *Blastocystis hominis*. *Parasitol Res* 83, 731-733.
- Zaman, V., Howe, J., Ng, M., Goh, T.K., 1999. Scanning electron microscopy of the surface coat of *Blastocystis hominis*. *Parasitol Res* 85, 974-976.
- Zambrano-Villa, S., Rosales-Borjas, D., Carrero, J.C., Ortiz-Ortiz, L., 2002. How protozoan parasites evade the immune response. *Trends Parasitol* 18, 272-278.
- Zhang, X., Qiao, J.Y., Zhou, X.J., Yao, F.R., Wei, Z.C., 2007. Morphology and reproductive mode of *Blastocystis hominis* in diarrhea and in vitro. *Parasitol Res* 101, 43-51.
- Zhou, X.B., Zhang, X., Qiao, J.Y., Cai, J., Cheng, S., Yuan, Y., Li, J., 2010. Encystation--survival of *Blastocystis hominis* in immunocompetent mice abdomen cavity. *Parasitol Res* 106, 1315-1320.
- Zierdt, C.H., Rude, W.S., Bull, B.S., 1967. Protozoan characteristics of *Blastocystis hominis*. *Am J Clin Pathol* 48, 495-501.
- Zierdt, C.H., 1973. Studies of *Blastocystis hominis*. *J Protozool* 20, 114-121.
- Zierdt, C.H., Williams, R.L., 1974. *Blastocystis hominis*: axenic cultivation. *Exp Parasitol* 36, 233-243.
- Zierdt, C.H., 1978. *Blastocystis hominis*, an intestinal protozoan parasite of man. Public Health Laboratory.
- Zierdt, C.H., 1986. Cytochrome-free mitochondria of an anaerobic protozoan--*Blastocystis hominis*. *J Protozool* 33, 67-69.
- Zierdt, C.H., 1988. *Blastocystis hominis*, a long-misunderstood intestinal parasite. *Parasitol Today* 4, 15-17.
- Zierdt, C.H., Donnelly, C.T., Muller, J., Constantopoulos, G., 1988. Biochemical and ultrastructural study of *Blastocystis hominis*. *J Clin Microbiol* 26, 965-970.
- Zierdt, C.H., 1991. *Blastocystis hominis* past and future. *Clin Microbiol Rev* 4, 61-79.
- Zucca, M., Savoia, D., 2011. Current developments in the therapy of protozoan infections. *Open Med Chem J* 5, 4-10.

Résumé

Blastocystis spp. est un Straménopile parasite anaérobie fréquemment rencontré dans le tractus gastro-intestinal de l'homme et de divers animaux. Ce parasite est associé à des troubles gastro-intestinaux aspécifiques, et semble impliqué dans des désordres fonctionnels tels que le syndrome de l'intestin irritable (IBS). Ce travail de thèse s'appuie sur le séquençage du génome de *Blastocystis* sp. ST7 réalisé en collaboration avec le Génoscope d'Evry, l'Université Nationale de Singapour, l'Institut Pasteur de Lille et l'Université de Provence. Ce génome est constitué d'un génome nucléaire de 18,8 Mpb pour 6020 gènes, et d'un génome mitochondrial de 29 kpb localisé dans des organites apparentés aux mitochondries. L'analyse de ce génome apporte des informations au niveau de l'évolution de ce microorganisme, de son adaptation à l'environnement intestinal et de ces facteurs de virulence potentiels. En effet, les analyses *in silico* de ce génome ont montré que *Blastocystis* sp. ST7 possède plusieurs gènes codant des protéines pouvant agir à l'interface entre l'hôte et le parasite et connues chez d'autres protozoaires pour être impliquées dans des phénomènes de pathogénie. Ce sont en particulier des PKS, des NRPS, et des hydrolases dont des protéases. D'autre part, des activités protéolytiques ont été mises en évidence expérimentalement dans les surnageants de culture du parasite. Deux protéases à cystéines (une cathepsine B et une légumaine) pouvant être impliquées dans la physiopathologie du parasite, ont été identifiées et caractérisées dans les surnageants, confirmant ainsi nos analyses *in silico*. Ce travail ouvre de nombreuses pistes intéressantes à explorer pour évaluer l'impact de ce parasite en santé humaine.

Mots-clefs : génome, analyse *in silico*, facteurs de virulence, protéases à cystéine.

Abstract

Blastocystis spp. is a highly prevalent anaerobic Stramenopile parasite found in the intestinal tract of humans and various animals. This parasite is associated with non specific intestinal disorders, and could be involved in functional disorders such as the irritable bowel syndrome (IBS). In this work, the *Blastocystis* sp. ST7 genome sequencing project was carried out in collaboration with the Génoscope of Evry, the National University of Singapore, the Pasteur Institute of Lille and the University of Provence. This genome consists in a nuclear genome of 18,8 Mpb encoding 6020 genes, and a mitochondria-like genome of 29 kpb localised in the mitochondrion-like organelles. The analysis of this genome brings information about the evolution of this micro-organism, its adaptation to the intestinal environment and its potential virulence factors. *Blastocystis* sp. ST7 was predicted to harbor several genes coding proteins that could act at the parasite-host interface, and that are known to be involved in the pathogeny of many protozoa. They are PKS, NRPS, and hydrolases among them proteases. In addition, proteolytic activities were highlighted in the parasite culture supernatants. Two cysteine proteases (a cathepsin B and a legumain) were identified and characterized from the supernatants and could play a role in the physiopathology of the parasite, that confirm our *in silico* analyses. This work opens new ways to evaluate the impact of this parasite in human health.

Key words : genome, in silico analyses, virulence factors, cysteine proteases